

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 53.406 Seconds
(without alignments)
825,861 Million cell updates/sec

Title: US-10-056-052A-4
Perfect score: 1732
Sequence: 1 MVAADAPAAAGTDTNQLTNV.....NGSGDGDIDKPVVPEQDPE 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1732	100.0	345	19	AAW31555
2	1727	99.7	933	21	AAW58435
3	1727	99.7	933	22	AAW69508
4	1723	99.5	936	18	AAW89801
5	1518	87.6	1021	22	AAU33975
6	1518	87.6	1021	22	AAU36951
7	1433	82.7	496	23	AAU75490
8	409	23.6	767	22	AAU34403
9	409	23.6	767	22	AAU37547
10	409	23.6	940	11	AAU07070

11	353	20.4	930	23	ABP40469	Staphylococcus epi
12	353	20.4	991	21	AAW83171	Cell wall protein
13	353	20.4	991	21	AAW70120	Staph. epidermidis
14	349	20.2	1092	19	AAW41602	Staphylococcus epi
15	346	20.0	978	22	AAU33960	Staphylococcus aur
16	346	20.0	1001	22	AAU37093	Staphylococcus aur
17	339	19.6	1018	22	AAU34301	Staphylococcus aur
18	339	19.6	1018	22	AAU37245	Staphylococcus aur
19	331	19.1	1018	9	AAW82115	Fibronectin bindin
20	331	19.1	1027	18	AAW89806	Staphylococcus aur
21	323.5	18.7	1166	20	AAW08643	S. aureus SdrC pro
22	312.5	18.0	918	20	AAW08640	S. aureus ClfB pro
23	282.5	16.3	251	18	AAW89804	Staphylococcus aur
24	257.5	14.9	1315	20	AAW08642	S. aureus SdrD pro
25	257.5	14.9	1349	22	AAU34402	Staphylococcus aur
26	257.5	14.9	1349	22	AAU37544	Staphylococcus aur
27	232.5	13.4	932	22	AAU34082	Staphylococcus aur
28	232.5	13.4	932	22	AAU36845	Staphylococcus aur
29	232	13.4	970	23	ABP27418	Streptococcus poly
30	210.5	12.2	841	22	AAU34283	Staphylococcus aur
31	210.5	12.2	841	22	AAU37158	Staphylococcus aur
32	210.5	12.2	930	20	AAW08641	S. aureus SdrC pro
33	204	11.8	1112	20	AAW08603	S. pyogenes SfrpP-
34	204	11.8	1161	23	AAE22273	Cell wall protein
35	201.5	11.6	1155	22	AAE22273	S. epidermidis ope
36	201.5	11.6	1802	21	AAW83170	Staphylococcus fibr
37	201.5	11.6	1802	21	AAW70119	Cell wall protein
38	157.5	9.1	462	23	ABW47305	Staph. epidermidis
39	140	8.1	2060	23	AAE20967	Listeria monocytog
40	138	8.0	1454	21	AAW56621	Staphylococcus lug
41	137	7.9	439	23	ABW47311	Neisseria meningit
42	136	7.9	559	18	AAW89803	Listeria monocytog
43	135	7.8	1983	23	ABW53455	Staphylococcus aur
44	132.5	7.7	1439	22	AAU27557	Lactococcus lactis
45	132.5	7.7	1457	21	AAW58592	Neisseria meningit
46	132.5	7.7	1457	21	AAW58592	N. meningitidis am
47	132.5	7.7	1457	21	AAW58592	N. meningitidis am
48	132.5	7.7	1457	22	AAU27556	Neisseria meningit
49	131.5	7.6	559	22	AAW82141	Neisseria meningit
50	127.5	7.4	999	23	AAW84052	S. epidermidis ope
51	127	7.3	793	23	ABW47318	S. pneumoniae deri
52	127	7.3	1057	21	AAW12725	Listeria monocytog
53	127	7.3	1058	23	AAW84097	Streptococcus pneu
54	126.5	7.3	571	23	ABW47317	S. pneumoniae deri
55	125.5	7.2	921	18	AAW22863	Listeria monocytog
56	125.5	7.2	921	21	AAW10626	Bacillus stearothe
57	125.5	7.2	1019	21	AAW12748	Streptococcus pneu
58	125.5	7.2	1019	21	AAW12752	Streptococcus pneu
59	125.5	7.2	1019	21	AAW12753	Streptococcus pneu
60	124.5	7.2	528	21	AAW12719	Streptococcus pneu
61	124.5	7.2	528	23	AAU84024	Truncated variant
62	124.5	7.2	568	21	AAW12731	Streptococcus pneu
63	124.5	7.2	568	23	AAW83999	Truncated variant
64	124.5	7.2	568	23	AAW84000	Truncated variant
65	124.5	7.2	568	23	AAW84008	Truncated variant
66	124.5	7.2	620	23	AAW84020	Truncated variant
67	124.5	7.2	626	23	AAW84018	Truncated variant
68	124.5	7.2	626	23	AAW84019	Truncated variant
69	124.5	7.2	632	23	AAW83998	Truncated variant
70	124.5	7.2	632	23	AAW84001	Truncated variant
71	124.5	7.2	632	23	AAW84002	Truncated variant
72	124.5	7.2	632	23	AAW84003	Truncated variant
73	124.5	7.2	632	23	AAW84004	Truncated variant
74	124.5	7.2	632	23	AAW84005	Truncated variant
75	124.5	7.2	632	23	AAW84006	Truncated variant
76	124.5	7.2	632	23	AAW84007	Truncated variant
77	124.5	7.2	632	23	AAW84009	Truncated variant
78	124.5	7.2	632	23	AAW84012	Truncated variant
79	124.5	7.2	632	23	AAW84013	Truncated variant
80	124.5	7.2	632	23	AAW84014	Truncated variant
81	124.5	7.2	632	23	AAW84015	Truncated variant
82	124.5	7.2	632	23	AAW84016	Truncated variant
83	124.5	7.2	632	23	AAW84017	Truncated variant

84 124.5 7.2 644 23 AAU83996 Truncated variant
 85 124.5 7.2 807 23 AAU83997 Truncated variant
 86 124.5 7.2 840 21 AAB12721 Streptococcus pneu
 87 124.5 7.2 840 23 AAU76151 Streptococcus pneu
 88 124.5 7.2 889 23 AAU84071 S. pneumoniae deri
 89 124.5 7.2 890 23 AAU84081 S. pneumoniae deri
 90 124.5 7.2 895 23 AAU84069 S. pneumoniae deri
 91 124.5 7.2 895 23 AAU84070 S. pneumoniae deri
 92 124.5 7.2 895 23 AAU84073 S. pneumoniae deri
 93 124.5 7.2 895 23 AAU84075 S. pneumoniae deri
 94 124.5 7.2 896 23 AAU84077 S. pneumoniae deri
 95 124.5 7.2 896 23 AAU84079 S. pneumoniae deri
 96 124.5 7.2 896 23 AAU84080 S. pneumoniae deri
 97 124.5 7.2 901 23 AAU84066 S. pneumoniae deri
 98 124.5 7.2 901 23 AAU84067 S. pneumoniae deri
 99 124.5 7.2 901 23 AAU84068 S. pneumoniae deri
 100 124.5 7.2 901 23 AAU84072 S. pneumoniae deri

ALIGNMENTS

RESULT 1

AAW31555
 ID AAW31555 standard; Protein: 345 AA.

XX
 AC AAW31555;

DT 21-MAY-1998 (first entry)

XX Fibronectin-binding MSCRAMM derivative pCF33.

DE Fibronectin; pCF33; collagen binding protein; sepsis; infection;
 KW microbial surface component recognising adhesive matrix molecule;
 KW MSCRAMM; adhesin; vaccine; immunisation; diagnosis; therapy;
 KW mastitis.

XX Staphylococcus aureus.

OS Staphylococcus aureus.

FH Key Location/Qualifiers
 FT Peptide 1..12
 FT /note= "vector pQE30-derived peptide"

XX WO9743314-A2.

XX 20-NOV-1997.

XX 14-MAY-1997; 97WO-US08210.

XX 16-MAY-1996; 96US-0017678.

XX (UABR-) UAB RES FOUND.
 PA (TEXA) UNIV TEXAS A & M SYSTEM.

XX Hook M, House-Pompeo K, Patti JM, Sthanam N, Symersky J;
 PI WPI; 1998-008801/01.

XX Antibody that interacts with collagen binding domain of

PT Staphylococcal cna gene product - useful to prevent bacterial sepsis
 PT in animal infected with Staphylococcus aureus

XX Disclosure; Page 91; 143pp; English.

XX This protein comprises Staphylococcus aureus fibronectin-binding
 CC microbial surface component recognising adhesive matrix molecule
 CC (MSCRAMM) derivative pCF33, plus a vector-derived N-terminal
 CC peptide. The invention relates to claimed nucleic acid sequences
 CC (see AAT93436-38) encoding S. aureus collagen binding protein (CBP)
 CC epitopes M17, M31 and M55 (see AAW31552-54) that confer protection
 CC against S. aureus infection. CBP protein and antigenic epitopes
 CC are contemplated for use in the treatment of pathological
 CC infections, especially to prevent bacterial adhesion to collagen.

CC The epitopes are also contemplated for use in the preparation of
 CC vaccines and as carrier proteins in vaccine formulations, as well
 CC as in the formulation of compositions for the prevention of S.
 CC aureus infection. pCF33 and pQD (see AAW31556) were used to raise
 CC anti-MSCRAMM polyclonal antibodies used in passive immunisation
 CC against bovine mastitis.

XX Sequence 345 AA;

Query Match 100.0%; Score 1732; DB 19; Length 345;
 Best Local Similarity 100.0%; Pred. No. 1.2e-116;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAADAPAAAGTDTNQLTNVTVGIDSGTTPVPHQAGYVKLVNFGFVSPNSAVKGDTEKITY 60
 |||||
 DB 13 MVAADAPAAAGTDTNQLTNVTVGIDSGTTPVPHQAGYVKLVNFGFVSPNSAVKGDTEKITY 72
 |||||

QY 61 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSGNVIVTFTDYVNTKDDVKATLTMPAYI 120
 |||||
 DB 73 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSGNVIVTFTDYVNTKDDVKATLTMPAYI 132
 |||||

QY 121 DPENVKKTGNVTLATGIGSTTANKTVLDVDEYKYGKYNLSIKGTTDQIDKTNNTYRQIY 180
 |||||
 DB 133 DPENVKKTGNVTLATGIGSTTANKTVLDVDEYKYGKYNLSIKGTTDQIDKTNNTYRQIY 192
 |||||

QY 181 VNPSGDNVIAPVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFVNPENFEDVT 240
 |||||
 DB 193 VNPSGDNVIAPVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFVNPENFEDVT 252
 |||||

QY 241 NSVNTTFPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGDLALRSTLYGYSNIWR 300
 |||||
 DB 253 NSVNTTFPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGDLALRSTLYGYSNIWR 312
 |||||

QY 301 SMSWDNEVAFNNGSGSGDGIDKPVVPEQDPE 331
 |||||

DB 313 SMSWDNEVAFNNGSGSGDGIDKPVVPEQDPE 343
 |||||

RESULT 2

AAV58435
 ID AAV58435 standard; Protein: 933 AA.

XX
 AC AAV58435;

DT 27-MAR-2000 (first entry)

XX Staphylococcus aureus fibrinogen binding ClfA protein.

KW ClfA; fibrinogen binding protein; bacterial colonisation;
 KW indwelling medical device; staphylococcal infection.

XX Staphylococcus aureus.

FH Key Location/Qualifiers
 FT Peptide 1..39
 FT /note= "Signal peptide"
 FT Region 40..559
 FT /note= "Region A"
 FT Region 332..550
 FT /note= "Fibrinogen-binding region"
 FT Region 560..867
 FT /note= "Region R"
 FT Region 896..900
 FT /note= "Gram positive wall-associated consensus motif"

XX US6008341-A.

XX 28-DEC-1999.

XX 22-AUG-1994; 94US-0293728.

XX 22-AUG-1994; 94US-0293728.

(QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

Foster TJ, McDevitt DL;

WPI: 2000-096389/08.

N-PSDB; AA25832.

Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus aureus, useful for treatment or prevention of infections

Disclosure; Fig 2A-1-4; 35pp; English.

This sequence represents the Staphylococcus aureus fibrinogen-binding protein, Clfa. Clfa is an important receptor involved in S. aureus colonisation of indwelling medical devices (e.g., catheters, artificial heart valves). Shortly after implantation, the surfaces of medical devices become coated with host plasma and matrix proteins such as fibrinogen and fibronectin, and there is considerable evidence to suggest that bacterial adherence to fibrinogen/fibrin is important in the initiation of device-related infection. The fibrinogen-binding region of Clfa is thought to reside between residues 332 and 550 in a region designated A. The protein also contains a repeated region (region R) comprising 154 repeats of the dipeptide Ser-Asp, and the C-terminus contains features present in surface proteins of other Gram positive bacteria that are responsible for anchoring the protein to the cell wall and cell membrane. Clfa, or its fragments, may be used to block S. aureus colonisation of wounds, to prevent adherence of S. aureus to indwelling medical devices, as vaccines to protect against S. aureus infection (e.g., mastitis in ruminants), to raise specific antibodies, and for diagnosis (by agglutination or immunosay). The specific antibodies are used for passive immunisation, to block infection of wounds or adhesion of S. aureus and for diagnosis. Nucleotides encoding Clfa and its fragments may be used as diagnostic probes.

Sequence 933 AA;

Query Match 99.7%; Score 1727; DB 21; Length 933;

Best Local Similarity 100.0%; Pred. No. 9.5e-116;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAADAPAGTDTITNQLTNVTGIDSGTTVYPHQAGYVKLVNYSFVSPNSAVKGDFTKITYP 61
DB 221 VAADAPAGTDTITNQLTNVTGIDSGTTVYPHQAGYVKLVNYSFVSPNSAVKGDFTKITYP 280
QY 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 121
DB 281 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 340
QY 122 PENVKKTGNVTLATGIGSTTANKTVLVVDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYV 181
DB 341 PENVKKTGNVTLATGIGSTTANKTVLVVDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYV 400
QY 182 NPSGDNVTAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVNPENEDVTN 241
DB 401 NPSGDNVTAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVNPENEDVTN 460
QY 242 SVNITFPNPNOYKVFENTPDDQITTPYIVVNGHIDPNSKGDALRLSTLYGYNLSNIWRS 301
DB 461 SVNITFPNPNOYKVFENTPDDQITTPYIVVNGHIDPNSKGDALRLSTLYGYNLSNIWRS 520
QY 302 MSWDNEVAFNNGSGGDGIDKPVVPEQDPE 331
DB 521 MSWDNEVAFNNGSGGDGIDKPVVPEQDPE 550

RESULT 3

AAB69508

ID AAB69508 standard; Protein: 933 AA.

XX AAB69508;

AC AAB69508;

XX 23-APR-2001 (first entry)

DT 23-APR-2001 (first entry)

XX

DE Staphylococcus aureus Clfa protein.

XX

KW Staphylococcus aureus; clfa; antibiotic; vaccine;

KW

XX fibrinogen binding protein; bacterial infection; mastitis.

XX

OS Staphylococcus aureus.

XX

PN US6177084-B1.

XX

PD 23-JAN-2001.

XX

PF 19-OCT-1999; 99US-0421868.

XX

PR 22-AUG-1994; 94US-0293728.

XX

PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

XX

PI Foster TJ, McDevitt DL;

XX

DR WPI: 2001-181608/18.

XX

DR N-PSDB; AAF58593.

XX

PT Novel Staphylococcus aureus fibrinogen binding protein useful as

PT

XX vaccine to prevent infection, promote wound healing, block adherence to

PT

XX indwelling medical devices and for diagnosing staphylococcus aureus

PT

PS infection

XX

PS Claim 5; Fig 2; 30pp; English.

XX

XX The present sequence is a novel Staphylococcus aureus fibrinogen

CC

CC binding protein. It is useful as a vaccine to protect against human

CC

CC and animal infections caused by S.aureus, such as against mastitis, to

CC

CC block S. aureus from colonising and infecting a wound, to block adherence

CC

CC of S.aureus to indwelling medical devices such as catheters, replacement

CC

CC heart valves and cardiac assist devices. The protein can be used to

CC

CC diagnose bacterial infections. Polyclonal and monoclonal antibodies

CC

CC raised against the fibrinogen binding protein are useful for passively

CC

CC immunising against infections caused by S.aureus, to prevent infection

CC

CC of a wound and to diagnose bacterial infections.

XX

SQ Sequence 933 AA;

XX

Query Match 99.7%; Score 1727; DB 22; Length 933;

XX

Best Local Similarity 100.0%; Pred. No. 9.5e-116;

XX

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 2 VAADAPAGTDTITNQLTNVTGIDSGTTVYPHQAGYVKLVNYSFVSPNSAVKGDFTKITYP 61

XX

DB 221 VAADAPAGTDTITNQLTNVTGIDSGTTVYPHQAGYVKLVNYSFVSPNSAVKGDFTKITYP 280

XX

QY 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 121

XX

DB 281 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 340


```
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
```

Sequence 1021 AA;

Query Match	87.68;	Score	1518;	DB	22;	Length	1021;
Best Local Similarity	87.3%;	Pred. No.	1.1e-100;				
Matches	288;	Conservative	17;	Mismatches	25;	Indels	0;
Gaps	0;						
Qy	2	VAADAPAAAGTDITNQLTNVTWVG	ISDGGTTVPVPHQAGYVKLN	YGFSVPNSAVKGD	TFKINVP	61	
Db	221	VAADAPAAAGTDITNQLTNVKT	ISDGGTTVPVPHQAGYVKLN	TGFSVPNSAVKGD	TFKINVP	280	
Qy	62	KELNLNGVTSTAKVPPIMAGDOVL	ANGVIDS	DGNVIYTF	TDYVNTKDDVKATL	MPAYID	121
Db	281	KELNLNGVTSTAKVPPIMVGDOVL	ANGVIDS	DGNVIYTF	TDYVNTKENVTANIT	MPAYID	340
Qy	122	PENVYKGTGNTLATIGIGSTTANK	TVLDVDEYKGYKFNLS	IKGTIDQIDKTN	TNTYRQTIYV	181	
Db	341	PENVYKGTGNTLTIGIGTNTASK	TVLDIYKGYGFHLS	IKGTIDQIDKTN	TNTYRQTIYV	400	
Qy	182	NPSGDNVIAPIVLTGNLKPNTDS	NALIDQONT	SIKVKYVDNAAD	ISESYFVNPFED	VTN	241
Db	401	NPSGDNVVLVLTGNLIPNTKSN	ALIIDAKNT	DIKRYVRD	NANDISESYVNP	SFEDVTN	460
Qy	242	SVNITFPNPQYKVEFTPDQIT	TPYVVVNGHIDPN	SKSGDLALR	SLYGYNSNI	IWRS	301
Db	461	QVRISFPNANQYKVEFTDDQIT	TPYVVVNGHIDP	ASTGDALR	STFYGYDSNF	IWRS	520
Qy	302	MSWDNEVAFNNGSGSGD	IKDPVVP	PEQDPE	331		
Db	521	MSWDNEVAFNNGSGSGD	IKDPVVP	PEQDPE	550		

RESULT 6
AAU36951
ID AAU36951 standard; Protein: 1021 AA.

AAU36951:

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #1121.

Antisense; prokaryotic cellular proliferation protein;
antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

WO200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206846P;
PR 26-MAY-2000; 2000US-207727P;

PR 23-OCT-2000; 2000US-242378P;
PR 27-NOV-2000; 2000US-253625P;

PR 22-DEC-2000; 2000US-237931P;
PR 16-FEB-2001; 2001US-269308P;

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX
FI, GAMMOCO KI, XU HH;

DR WPI; 2001-011493770.
DR N-PSDB; AAS54810.

xx New polynucleotides for the identification and development of
PT

FI antibiotics, comprise sequences of antisense nucleic acids -

XX PS XX CC

Example 3; Seq ID No 12544; 511pp; English.

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 1021 AA;

Query Match	87.6%	Score 1518;	DB 22;	Length 1021;
1st Local Similarity	87.3%;	Pred. NO. 1.1e-100;		
Matches 288;	Conservative	17;	Mismatches 25;	Indels 0;
Gaps 0				

Qy	2 VAADAPAACTDITNQLTNVTGIDSGTTVYPHQAGVVKLVNGFSVPNSAVKGDTFKITVP 61 :
Db	221 VAADAPAACTDITNQLTDYKVKTDSGTTVYPHQAGVVKLVNGFSVPNSAVKGDTFKITVP 280

Qy	62	KEINLNGVSTAKVPPIMAGDQVLIANGVIDSQGNVIYFTFYVNTKDDVKATLTM	PAYID	121
Dh	281	KEINLNGVSTAKVDDTMYGQDVVANGVTDSCNVITVETDGVVGVGVNANTM	PAYID	240

Qy 122 PENVKKTGNVTLATGIGSTTANKTTLVDVEKYKGFYNLSIKGTIDQIDKTNNTYRQTIYV 181

QY 182 NP SGDNV IAPVL TGNL KPN TDSNAL IDQQNTS I K Y K V D N A A D L S E S Y F V N P E N F E D V T N 241

Qy 242 SVNITFPNPNQYKVEFNTDDQITPYIVVVNGHIDPNKGDALRSTLYGVNSNTIWR\$ 301

50
 401 QVRLSFNFANQINVEFFIDDDQIIIFIVVNGRHDFASIGDLAKLSIFIGIDSNELWKS 322
 Qy 302 MSWDNEVAFNNGSGSGDIDKPVVPEQPDE 331
 |||||

DB 521 MSWDNEVAFNNGSGSGDIDKPVWPPEQDE 550

RESULT 7

AAU75490
ID AAU75490 standard: 496 AA

XX
AC AAI75490:XX
DT 23-APR-2002 (first entry)

XX
DE
S. aureus antigenic protein associated protein #10.

XX Antigenic protein: vaccine; SPREX; antibacterial;
XX antiinflammatory; dermatological; antiulcer; tuberculostatic;
KW immunosuppressive; septicaemia; food poisoning; skin disorders;
KW peritonitis; endocarditis; tuberculous; blood infection; sepsis;
KW meningitis; pneumonia; stomach ulcer; gonorrhea; necrotising fasciitis;
KW impetigo; Lyme's disease; gastro-enteritis; dysentery; shigellosis.
XX
XX Staphylococcus aureus.
OS
XX

Db 155 VVEETKATGTDVTK-----VEVEEGSEIVGHKQDTNVNPNHAEVTLKYKWKFGEGIK 209
 QY 52 KGDTEKITVPKELNLNGVTSTAKVPPIMAGD-OVLANGVIDSDGNVIYFTTDDYVNTKDDV 110
 Db 210 AGDYDFDLSDNVETHGISTLRKPEIKSTDCQVWATGEIIGERKVRVTFEYVQEKD 269
 QY 111 KATLMPAYIDPENVKGTGNVTLATGISTTANKTVLVY-----EYKGFKNLSIKGTI 165
 Db 270 TAEISLNLFDPTTVTQKGNQNVKLGSETTVSKIFNIQYLGVRDNGM---VTANGRI 325
 QY 166 DOIDKNTNNTYQTIYVNPNSGDNVIAPVLTGNL-----KPTDSNALIDQNTSIKVKYVDN 221
 Db 326 DTLNKVDGKSFHAYMKPNQSLSSVTGTQVTKGNKPGVN-----NPTVKYKHI 377
 QY 222 AADLSESYFVNPN---FEDVTNSVNIITFPNPNQYKVEFNTPDDOITTPYIVVNGHIDP 278
 Db 378 SDDLAESVYAKLDDVSKFEDVTDNMSLDFDTNGGYSLNFNLDQ--SKNYVIRYEGYIDS 435
 QY 279 NSKGDLALRSTLYGYNSNIIRSMNSWDNEVAFNNGSGDGDIDK---PWV 325
 436 NA-SNLEFQTHLFGYNYNYTSLNLTWKNGVAFYSNNAQDGDKLKEPII 484

RESULT 9

AAU37547
 ID AAU37547 standard; Protein; 767 AA.

XX AAU37547;

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #1717.

XX Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 22-DEC-2000; 2000US-257931P.

XX 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX N-PSDB; AAS55406.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 13140; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the

XX genes, their use in the discovery of novel antibiotics, the essential

CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 767 AA;

Query Match 23.6%; Score 409; DB 22; Length 767;

Best Local Similarity 28.9%; Pred No. 4.2e-21;

Matches 101; Conservative 72; Mismatches 131; Indels 46; Gaps 11;

QY 2 VAADAPAAAGTDTITNQLTNVTGIDSGT-----TVYPHQAQYVKNLNGFVSVPNSAV 51

Db 155 VVEETKATGTDVTK-----VEVEEGSEIVGHKQDTNVNPNHAEVTLKYKWKFGEGIK 209

QY 52 KGDTEKITVPKELNLNGVTSTAKVPPIMAGD-OVLANGVIDSDGNVIYFTTDDYVNTKDDV 110

Db 210 AGDYDFDLSDNVETHGISTLRKPEIKSTDCQVWATGEIIGERKVRVTFEYVQEKD 269

QY 111 KATLMPAYIDPENVKGTGNVTLATGISTTANKTVLVY-----EYKGFKNLSIKGTI 165

Db 270 TAEISLNLFDPTTVTQKGNQNVKLGSETTVSKIFNIQYLGVRDNGM---VTANGRI 325

QY 166 DOIDKNTNNTYQTIYVNPNSGDNVIAPVLTGNL-----KPTDSNALIDQNTSIKVKYVDN 221

Db 326 DTLNKVDGKSFHAYMKPNQSLSSVTGTQVTKGNKPGVN-----NPTVKYKHI 377

QY 222 AADLSESYFVNPN---FEDVTNSVNIITFPNPNQYKVEFNTPDDOITTPYIVVNGHIDP 278

Db 378 SDDLAESVYAKLDDVSKFEDVTDNMSLDFDTNGGYSLNFNLDQ--SKNYVIRYEGYIDS 435

QY 279 NSKGDLALRSTLYGYNSNIIRSMNSWDNEVAFNNGSGDGDIDK---PWV 325

Db 436 NA-SNLEFQTHLFGYNYNYTSLNLTWKNGVAFYSNNAQDGDKLKEPII 484

RESULT 10

AAU37547

ID AAR07070 standard; protein; 940 AA.

XX AAR07070;

XX 21-FEB-1991 (first entry)

XX Fibronectin-binding protein encoded by gene 2, fnbb.

XX Wound healing; Immunisation; mastitis; fnbA; ss.

XX Staphylococcus aureus.

XX Key Location/Qualifiers

XX Peptide 1..36

XX /label=signal peptide

XX EP397633-A.

XX 14-NOV-1990.

XX 04-MAY-1990; 90EP-0850166.

XX 14-NOV-1990; 90SE-0001687.

XX (ALFA) ALFA-LAVAL AGRI INT.

XX Hook M, Jonsson KL, Lindberg KM, Signas LC;

XX

DR WPI: 1990-343563/46.
DR N-PSDB; AAQ06550.

XX Fibronectin binding protein gene - isolated from *S. aureus* and used
PT for producing protein for wound healing, immunisation or diagnosis.
XX

PS Claim 11; Page 17; 24pp; English.

XX The sequence is deduced from a gene, designated gene 2 (fnbB),
CC localised downstream of that disclosed in SE8702272 (gene 1; fnbA)
CC encoding a fibronectin binding protein, FnBPA in *S. aureus*. Gene 2
CC encodes fnbPB. The two proteins show considerable homology, esp.
CC in the signal sequence (95%). Downstream is a stretch of 66 AAs
CC with 75% homology to the same stretch in FnBPA. The following 44
CC AAs have only 40% homology and the B repeats of FnBPA are absent.
CC However the rest of the peptide (394 AAs) is nearly identical to
CC FnBPA, the main difference being a deletion of 14 AAs. This highly
CC homologous region contains the same repeat (D1-D4 and W1-5) found
CC in FnBPA with the exception that W1 is missing. The WC region and
CC the hydrophobic M domain as well as the C-terminal end are conserved.

Sequence 940 AA;

Query Match 23.6%; Score 409; DB 11; Length 940;

Best Local Similarity 28.9%; Pred. No. 5.4e-21;

Matches 101; Conservative 72; Mismatches 131; Indels 46; Gaps 11;

QY 2 VAADAPAAAGTDTTNTLTNTVTVGIDSGT-----TVYPHOAGYVKNLYGFSVPNSAV 51
DB 155 VVEETKATGTDVTKN-----VEVEGSEIVGHKQDTNVNPNNAERVTLKWKKEGEGIK 209
QY 52 KGDTEKTIPTVKELNLTNGVTSTAKVPPIMAGD-QVLANGVIDSDGNVIYFTDYVNTKDDV 110
DB 210 AGDYFDTLSDNVETHGISTLRKPEIKSTKDGQVWATGEIIGERKVRVTFEYVQEKDL 269
QY 111 KATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVLDY-----EKYKGFYVLSIKGTI 165
DB 270 TAEKSLNLFIDTPTVQKGNQVNEVKLGETTVSKIFNIQYLGGRVDRNG---VTANGRI 325
QY 166 DQIDKTNNTYRTIYVNPVSGDNVIAPIVLTGNL---KPTDSNALIDQNTSIKVKYVDN .221
DB 326 DTLAKVDGKFSHFAYKPNKQSLSVTVGTQVTKGNKPGVN-----NPTVKYKHIG 377
QY 222 AADLESYEVNPN---FEDVTSNINITFPNPNQYKVEFNTPDQITTPYIVVNGHIDP 278
DB 378 SDDLAESVYAKLDDVSKEDVTDNMSLDFDTNGGYSLNFNLDQ--SKNYIKYEGYVDS 435
QY 279 NSKGLDALRLSTLYGYNLSNIWRMSWMDNEVAFNNGSGSGDGIDK---PVV 325
436 NA-SNLEFQTHLFGYNYNYTSLNLTWKNGVAFYSNNAOGDGKDKLKEPII 484

RESULT 11

ABP40469

ID ABP40469 standard; Protein; 930 AA.

XX ABP40469;

XX

24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5314.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy.

XX Staphylococcus epidermidis.

PN US6380370-B1.

XX

PD 30-APR-2002.

XX

PF 13-AUG-1998; 98US-0134001.

XX

PR 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX

PI Doucette-Stamm LA, Bush D;

XX

DR WPI: 2002-381255/41.

DR N-PSDB; ABN93014.

XX Novel isolated nucleic acid encoding a *Staphylococcus epidermidis*
PT polypeptide, useful for diagnosing and treating bacterial infections
XX
PS Disclosure; SEQ ID 5314; 267pp; English.

CC ABN90538 to ABN93374 represent *Staphylococcus epidermidis* open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The *S. epidermidis* sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly *S. epidermidis* infections. The sequences can be used to
CC screen for compounds able to interfere with the *S. epidermidis* life
CC cycle or inhibit *S. epidermidis* infection.

CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.

SQ Sequence 930 AA;

Query Match 20.4%; Score 353; DB 23; Length 930;

Best Local Similarity 30.6%; Pred. No. 5.7e-17;

Matches 101; Conservative 60; Mismatches 139; Indels 30; Gaps 14;

QY 13 ITNQLTNTVTVGI-DSGTTVYPHOAGYVKNLYGFSVPNSAVKCDTEKTIPTVKELNLTNGVTS 71.
DB 284 VTQD--SITEGDDSDGIIKAHAENLIYDVTFEVDDKVKSGDPTVVIDKNTVPDSDLTD 341
QY 72 TAKVPPIMAGD-QVLANGVID-SDGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTG 129
DB 342 SFAIPKIDNKGSEIATCTYDNTNKGIIYFTDYVYKVENIKAHKLKTSYIDKSKVPNN 401
QY 130 ---NVTLATGISTTANKTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYRQIYVNPVSGD 186
DB 402 TKLDEYKATLSS--VNKTITVEYKPNENRTANLQSMFTNIDTKNHTVEQTIYNPL-- 457
QY 187 NVIAPVLGTLKPNTPDSNALIDQNTSKYKVDNAADLSYSFV-NPENFEDVTNSVNI 245
DB 458 RYSAKETNVNISGNGDEGSTIIDSTIIKVKVGNQNLPSNRIYDSEYEDVTNDYYA 517
QY 246 TFPNPNQYKVEFNTPDQITTPYIVVNGHIDPNKSGD-----LALRSTLYGYNLSNI 298
DB 518 QLGNNDVNIINFG----NIDSPYIIKVKSYDPN-KDDYTTTIOQTVTWQTINEYTG-- 570
QY 299 WRMSWMDNEVAFNNGSGSGDGIDKPVVPEQ 328
DB 571 FRTASYDNTIAFSTSSGQCG-DLP--PEK 597

RESULT 12

AA83171

ID AA83171 standard; Protein; 991 AA.

XX AA83171;

XX

24-JUL-2000 (first entry)

DE Cell wall protein SdrG.

XX SdrF; SdrG; SdrH; coagulase negative; *Staphylococcus*; septicemia;
KW osteomyelitis; endocarditis; immune response; vaccine; graft;
KW stent; intravenous catheter; heart valve; cardiac.

OS *Staphylococcus* sp.

CC The present sequence is a serine-aspartate repeat region
 CC protein, SdrG from Staphylococcus epidermidis. The Sdr protein is
 CC useful in vaccine preparation in combination with specific
 CC bacterial binding proteins. These vaccines can be used to treat a broad
 CC spectrum of bacterial infections, including those arising from both
 CC coagulase-positive and coagulase-negative bacteria.

XX Sequence 991 AA;

Query Match 20.4%; Score 353; DB 21; Length 991;
 Best Local Similarity 30.6%; Pred. No. 6.1e-17;
 Matches 101; Conservative 60; Mismatches 139; Indels 30; Gaps 14;

QY 13 ITNQLTNTVTGI--DSGTTVPYHQGVYKLVNYSFVPSNAVKGDTFKITVPKELNLNGVTS 71
 Db 317 VTQD--SITEGYDDSGIIKAHDAENLIYDVTFEVDKVKSGDTWTVDKNTVPSDLTD 374

QY 72 TAKVPPIMAGD-QVLANGVID-SGNNVIYTFDYVNTKDDVKATLTPAYIDPENVKKTG 129
 Db 375 SPAIPKIDNSGELIATGYDNTNKGITVTFDYVDKVENIKAHKLKLTYSIDKSKVPNNN 434

QY 130 ---NVTLATGICSTTANKTVLDYKYGKFNLSKGTIDQDKNTNRYQIYVNPSSD 186
 Db 435 TKLDVEYKTLSS--VNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYNPL-- 490

QY 187 NVIAPVLTCNLKPNNTDSNALIDQNTSIRKVKVDNAADLSEYFV-NPENFEDVTNSVNI 245
 Db 491 RYSAKETNVNISNGDEGSTIIDSTIIKVKVGNQNLPSNRIDYSEYEDVNDYDA 550

QY 246 TFPNPQYKVFNTPDQITTPYIVVNGHIDPNSKGD-----LALRSTLYGNSNI 298
 Db 551 QLGNNNDVNIQF---NIDSPYIIKISKYDPN-KDDYTTTQQVTVMQTINEVTGE-- 603

QY 299 WRSMSWDNEVAFNNGSGGSDGIDKPVVPEQ 328
 Db 604 FRTASYDNTIAFSTSSGQGG-DLP--PEK 630

RESULT 14

ID AAW41602
 XX AAW41602 standard; Protein; 1092 AA.

AC AAW41602;

DT 22-JUN-1998 (first entry)

DE Staphylococcus epidermidis fibrinogen binding protein FIG.

XX Fibrinogen binding protein; FIG; aggregation; infection;
 coagulase-negative Staphylococcus; therapy; diagnosis;
 immunisation; immunogen; vaccine.

XX Staphylococcus epidermidis strain HB.

XX Key Location/Qualifiers

FT Peptide 1..51

FT Protein /label= Sig_peptide

FT Protein /label= Mat_protein

FT Region 52..824

FT /note= "non-repetitive region, harbours
 fibrinogen binding activity"

FT Region 825..1040

FT /note= "Asp-Ser dipeptide repeat region"

FT Region 1053..1057

FT /note= "cell wall anchoring motif"

XX WO9748727-A1.

XX 24-DEC-1997.

XX 18-JUN-1997; 97WO-SE01091.

XX

PR 20-JUN-1996; 96SE-0002496.

XX (FLOC/) FLOCK J.

PA (FRYK/) FRYKBERG L.

PA (GUSS/) GUSS B.

PA (LIND/) LINDBERG M.

PA (NILS/) NILSSON M.

XX

PI Flock J, Frykberg L, Guss B, Lindberg M, Nilsson M;

XX WPI: 1998-063079/06.

DR N-PSDB; AAV04279.

XX

PT Fibrinogen-binding protein from coagulase-negative Staphylococcus -
 PT used for prevention, treatment and diagnosis of Staphylococcus
 PT infection

XX

PS Example 3; Fig 6; 45pp; English.

XX The protein comprises the fibrinogen binding protein (FIG) of

CC coagulase-negative Staphylococcus epidermidis HB. Its amino

CC acid sequence was deduced from the isolated fig gene (see AAV04279).

CC The closest known analogue of FIG is the clumping factor of

CC S. aureus which also binds fibrinogen and promotes bacterial

CC aggregation in serum. Recombinant FIG polypeptides can be

CC expressed in host cells. They are used as immunogens, particularly

CC in vaccines (which may be expressed in vivo) to protect humans and

CC animals against coagulase-negative Staphylococcus infection.

CC Antibodies raised against FIG can be used for passive immunisation.

CC They block the adherence of bacteria) and for diagnosis.

XX Sequence 1092 AA;

QY Query Match 20.2%; Score 349; DB 19; Length 1092;

Best Local Similarity 30.3%; Pred. No. 1.4e-16;

Matches 100; Conservative 61; Mismatches 139; Indels 30; Gaps 14;

QY 13 ITNQLTNTVTGI--DSGTTVPYHQGVYKLVNYSFVPSNAVKGDTFKITVPKELNLNGVTS 71

Db 287 VTQD--SITEGYDDSGEVIKAHDAENLIYDVTFEVDKVKSGDTWTVDKNTVPSDLTD 344

QY 72 TAKVPPIMAGD-QVLANGVIDS-DGNNVIYTFDYVNTKDDVKATLTPAYIDPENVKKTG 129

Db 345 SFTPIKIDNSGELIATGYDNTNKGITVTFDYVDKVENIKAHKLKLTYSIDKSKVPNNN 404

QY 130 ---NVTLATGICSTTANKTVLDYKYGKFNLSKGTIDQDKNTNRYQIYVNPSSD 186

Db 405 TKLDVEYKTLSS--VNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYNPL-- 460

QY 187 NVIAPVLTCNLKPNNTDSNALIDQNTSIRKVKVDNAADLSEYFV-NPENFEDVTNSVNI 245

Db 461 RYSAKETNVNISNGDEGSTIIDSTIIKVKVGNQNLPSNRIDYSEYEDVNDYDA 520

QY 246 TFPNPQYKVFNTPDQITTPYIVVNGHIDPNSKGD-----LALRSTLYGNSNI 298

Db 521 QLGNNNDVNIQF---NIDSPYIIKISKYDPN-KDDYTTTQQVTVMQTINEVTGE-- 573

QY 299 WRSMSWDNEVAFNNGSGGSDGIDKPVVPEQ 328

Db 574 FRTASYDNTIAFSTSSGQGG-DLP--PEK 600

XX

XX RESULT 15

XX AAU33960

ID AAU33960 standard; Protein; 978 AA.

XX AAU33960;

XX

XX 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #236.

XX Antisense; prokaryotic cellular proliferation protein;

XX

Query Match 20.0%; Score 346; DB 22; Length 1001;
 Best Local Similarity 27.2%; Pred. No. 2e-16;
 Matches 91; Conservative 69; Mismatches 148; Indels 26; Gaps 10;

QY 3 AADAPA-AGTDTITNLTNTVTVGIDS--GTTVYPHOAGVYKLVNYSFVNSAVKGDFTFKIT 59
 DB 186 ASDAKVETGTDVTSKVTVEGSIAPQGNKVEPHAGORVLYKLKFEKGLHKGDFDFT 245
 QY 60 VPKEINLNGVSTAKVPPIMAGDOVLANGVIDSDGNVIYTTDXYNTKDDVKATLTWPAY 119
 DB 246 LSNVNTVGVSTARKVPEIKGVSVMATGQLGNGKIRYTTFTDIDYKVNVTADLEINLF 305
 QY 120 IDPENVKKTNGVTLATGIGSTTANKTVLDVYEKGYKFNLSIKGTIDQIDKTNNTYRGTI 179
 DB 306 IDPKTVQSQOQTITSTLNDRETKNTLPIEYNGVSNYSYANVNGSIETFDKGNRFTHVA 365
 QY 180 YVNP-SGDNVITAPVLTGKLNKPTDSDNALIDQNTSIKVKY-DNAADLSESYFVNPN-- 235
 DB 366 YIKPQNGHKSQSVSITGTLTQGSKA---DGKAPTVMKYEVLDKANELPQSVYANVSDSS 421
 236 -FEDVTSV--NITFPNPNQYKVEFNTPDQDITTPYIVVNGHIDPNKSGDLALRSTLYG 292
 422 MFKDVTETMOKLKVNNNGKLDI----EKLEKSYVHYDGEYLSGS-DQVNFRTMFG 476

QY 293 YNS-----NIWRMSMDNEVAFNNGSGSDG 319
 DB 477 YPEQGYKYVYTHLGYKLTWDNGLVLYSNKAKGDG 510

RESULT 17
 ID AAU34301 standard; Protein; 1018 AA.
 AC AAU34301;
 XX
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Staphylococcus aureus cellular proliferation protein #577.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Staphylococcus aureus.
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001WO-US09180.
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 XX WPI; 2001-611495/70.
 DR N-PSDB; AAS52160.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics; comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 5797; 511pp; English.
 CC
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the

genes, their use in the discovery of novel antibiotics, the essential
 genes themselves and the encoded proteins. The prokaryotes used are
 Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 invention is also useful for the identification of potential new targets
 for antibiotic development. The antisense nucleic acids can also be used
 to identify proteins used in proliferation, to express these proteins,
 and to obtain antibodies capable of binding to the expressed proteins.
 The proteins can be used to screen compounds in rational drug discovery
 programmes. The antisense nucleic acid sequence is also useful to screen
 for homologous nucleic acids which are required for cell proliferation in
 a wide variety of organisms. The present sequence represents an
 essential prokaryotic cellular proliferation protein.
 Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic
 format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.

Sequence 1018 AA;
 SQ
 Query Match 19.6%; Score 339; DB 22; Length 1018;
 Best Local Similarity 24.9%; Pred. No. 6.5e-16;
 Matches 83; Conservative 77; Mismatches 145; Indels 28; Gaps 11;

QY 10 GTDITNLTNTVTVGIDSG---TTVYPHOAGVYKLVNYSFVNSAVKGDFTFKITVPKELN 65
 DB 194 GTDVTSKVT-VEIGSIEGHNTNKVEPHAGORAVLYKLKFEKGLHOGDYDFDTLSNNVN 252
 QY 66 LNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYTTDXYNTKDDVKATLTWPAYIDPNV 125
 DB 253 THGVSTARKVPEIKGVSVMATGVELEGKRYTFTNDIEDKVDVTALEINLFIDPKTV 312
 QY 126 KKTGNVTLATGIGSTTANKTVLDVY-EKGYKFNLSIKGTIDQIDKTNNTYRQIYVNP 184
 DB 313 QTNGNQITITLNEEQTSKELDVYKDGIGNY-ANLNGSIETFNKANNRSHVAFKPN 371
 QY 185 GDNVIAPVLTGKLNKPTDSDNALIDQNTSIKVKY-VDNAADLSESYFVN---PENFEDVT 240
 DB 372 NGKTTSTVTVTGTLMKGSNQNG---NQPKVRIFEVLGNNEEDIAKSVYANTDTSKFREVT 427
 QY 241 NSV--NITFPNPNQYKVEFNTPDQDITTPYIVVNGHIDPNKSGDLALRSTLYGNSII 298
 DB 428 SNMSGNLNLQNGSYSLSNI----ENLDKTYVYHYDGEY-LNGTDEVDFTQMVGHPEOLY 482
 QY 299 -----WRMSMDNEVAFNNGSGSDGIDKPVV 325
 DB 483 KYYYDRGYTLTWDNGLVLYSNKANGKNGKNGPII 515

RESULT 18
 AAU37245
 ID AAU37245 standard; Protein; 1018 AA.
 XX
 AC AAU37245;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Staphylococcus aureus cellular proliferation protein #1415.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Staphylococcus aureus.
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001WO-US09180.
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI: 2001-611495/70.
 DR N-PSDB; AAS55104.
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 12838; 511pp; English.
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1018 AA;
 Query Match 19.6%; Score 339; DB 22; Length 1018;
 Best Local Similarity 24.9%; Pred. No. 6.5e-16;
 Matches 83; Conservative 77; Mismatches 145; Indels 28; Gaps 11;
 QY 10 GTDITNQLTNTVTVGIDSG-----TTVYPHQAGYVVKLYGFSVPNSAVKGDTEKITVPKELN 65
 DB 194 GTDVTSKYT-VEIGSIEGHNTNKNVEPHAGQRAVLKYLKFKENGHLHGDYDFDTLSNNVN 252
 DB 66 LNGVTSTAKVPPIMAGDOVLANGVITSDGNVIYTFDVTYNTKDDVKATLTMPAYIDPENY 125
 DB 253 THGVSTARKVPEIKNGSVVMATGEVLEGGKIRYTFETNDIODEKVDVTAELNLFDPKTV 312
 QY 136 KKTGNVTATGIGSTANKTVLVDY-EKYGFYNSIKGTIDQIDKNTTYRQTIYVNP 184
 DB 313 QTNGNQITITSLNEBQTSKELDVKYKDGIGNY-ANLNGSIETFNKANNRFSHVAFIKPN 371
 QY 185 GDNVIAPVLTGNLKNPNTSDNALIDQONTSIKVKY-VDNAADLSSEYFVN---PENFEDVT 240
 DB 372 NGKTSVTVTGTLMKGSNQNG-----NPKVRIFEYVLGNNEIDAKSVYANTTDTTSKEV 427
 QY 241 NSV---NITFPPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGNSNII 298
 DB 428 SNMGSNLLQNGSYSLNI-----ENLDKTYVVHYDGEY-LNGTDEVDFRTQMVGHPEQLY 482
 QY 299 -----WRSMGWDNEVAFNNGSGSGDGIDKPVV 325
 DB 483 KYYYDRGYTLTWDNGLVLYSNKANGKNGNP 515
 RESULT 19
 AAP82115
 ID AAP82115 standard; protein; 1018 AA.
 XX

AC AAP82115;
 XX 05-JAN-1990 (first entry)
 DT Fibronectin binding protein.
 DE Fibronectin binding protein.
 XX Fibronectin binding protein; S.aureus; immunisation;
 KW mastitis; ruminants; wound infection; diagnosis.
 OS Staphylococcus aureus.
 XX EP294349-A.
 PN 07-DEC-1988.
 PD 30-MAY-1988; 88EP-0850188.
 PF 01-JUN-1987; 87SE-0002272.
 PR (ALFA) ALFA-LAVAL AGRI INTERNATIONAL AB.
 PA Hook M, Lindberg MK, Signas LC, Wadstrom TM, Froman G;
 PI WPI: 1988-347978/49.
 XX P-PSDB; AAP82115.
 DR Hybrid DNA encoding staphylococcus aureus fibronectin binding protein -
 PT useful for immunisation and topical application to prevent
 PT staphylococcal infections.
 XX Disclosure; Fig. 8A:1-8A:2; 23pp; English.
 PS The S.aureus fibronectin binding protein may be included in a hybrid
 CC protein. The protein may be used to immunise ruminants against
 CC staphylococcal mastitis, pref. when given at 0.5-5 microg/kg, using 3
 CC doses at 1-3 week intervals; and for topical applicn. to prevent wound
 CC infection, pref. using isotonic saline soln. of concn. 25-250 microg/mg.
 CC The DNA encoding the binding protein and the hybrid protein may be
 CC immobilised on a carrier and used to diagnose staphylococcal infections.
 XX
 SQ Sequence 1018 AA;
 Query Match 19.1%; Score 331; DB 9; Length 1018;
 Best Local Similarity 24.6%; Pred. No. 2.4e-15;
 Matches 82; Conservative 77; Mismatches 146; Indels 28; Gaps 11;
 QY 10 GTDITNQLTNTVTVGIDSG-----TTVYPHQAGYVVKLYGFSVPNSAVKGDTEKITVPKELN 65
 DB 194 GTDVTSKYT-VEIGSIEGHNTNKNVEPHAGQRAVLKYLKFKENGHLHGDYDFDTLSNNVN 252
 QY 66 LNGVTSTAKVPPIMAGDOVLANGVITSDGNVIYTFDVTYNTKDDVKATLTMPAYIDPENY 125
 DB 253 THGVSTARKVPEIKNGSVVMATGEVLEGGKIRYTFETNDIODEKVDVTAELNLFDPKTV 312
 QY 136 KKTGNVTATGIGSTANKTVLVDY-EKYGFYNSIKGTIDQIDKNTTYRQTIYVNP 184
 DB 313 QTNGNQITITSLNEBQTSKELDVKYKDGIGNY-ANLNGSIETFNKANNRFSHVAFIKPN 371
 QY 185 GDNVIAPVLTGNLKNPNTSDNALIDQONTSIKVKY-VDNAADLSSEYFVN---PENFEDVT 240
 DB 372 NGKTSVTVTGTLMKGSNQNG-----NPKVRIFEYVLGNNEIDAKSVYANTTDTTSKEV 427
 QY 241 NSV---NITFPPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGNSNII 298
 DB 428 SNMGSNLLQNGSYSLNI-----ENLDKTYVVHYDGEY-LNGTDEVDFRTQMVGHPEQLY 482
 QY 299 -----WRSMGWDNEVAFNNGSGSGDGIDKPVV 325
 DB 483 KYYYDRGYTLTWDNGLVLYSNKANGKNGNP 515
 RESULT 20
 AAP82115

ID	AAW89806	standard; Protein; 1027 AA.
XX		
AC	AAW89806;	
XX		
DT	16-MAR-1999	(first entry)
XX		
DE	Staphylococcus aureus	protein SEQ ID #5254.
XX		
KW	Computer readable medium; vaccine; S.aureus	infection; immunodetection;
KW	cellulitis; eyelid	infection; food poisoning; osteomyelitis; therapy;
KW	skin infection; surgical	wound infection; scalded skin syndrome;
KW	toxic shock	syndrome.
XX		
OS	Staphylococcus aureus.	
XX		
PN	EP786519-A2.	
XX		
PD	30-JUL-1997.	
XX		
07-JAN-1997;	97EP-0100117.	
XX		
05-JAN-1996;	96US-0009861.	
XX		
(HUMA-)	HUMAN GENOME SCI INC.	
XX		
PA	Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;	
PI	Rosen CA;	
PI		
XX		
DR	WPI; 1997-374922/35.	
XX		
XX	Polynucleotide(s) and proteins derived from Staphylococcus aureus	
PT	stored on computer readable medium and used in the production of	
PT	anti-S.aureus vaccines	
XX		
PS	Claim 23; Page 3263-3267; 3271pp; English.	
XX		
XX	This sequence represents a Staphylococcus aureus protein sequence of the invention. The DNA sequences encoding the S.aureus proteins are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.	
CC	Homology searches using the S.aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained.	
CC	Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against S.aureus infection. The polypeptides can also be used in a kit for the immunodetection of S.aureus in a sample. S.aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the 5191 S.aureus DNA sequences contained on the computer readable medium.	
CC		
XX	Sequence	1027 AA;
XX		
Query Match	19.1%;	Score 331; DB 18; Length 1027;
Best Local Similarity	24.6%;	Pred. No. 2.5e-15;
Matches	82; Conservative	77; Mismatches 146; Indels 28; Gaps 11;
QY	10	GTDTNLTNTVTVGIDSG----TTVYPHOAGVVKLVNFGFSPNSAVKGDTEKITVPKELN 65
DB	203	GTDTVTSKVT-VEIGSIEGHNTNKNVEPRAGORAVLKYKLFENGHLHQGDYDFDTLSNNVN 261
QY	66	LVNGTSTAKVPPIMAGDQVLANGVSDSGNIYTFDTYVNTKDDVKATLTMPAYIDPNV 125
DB	262	THGVSTARKVPEIKNSVVMATGEVLEGGKIRYFTFTNDIEDKVDVDALEINLFDIPKTV 321
QY	126	KKTGNVTATIGSTANKTVLDVY-EKYGFXNLSIKGTIDQIDKTNNTVQTIIYVNS 184
DB	322	QTNGNOTITLTINEEQTSKELDVKYDKIGNYY-ANLNGSIETETNKANNRFSHVAFIKPN 380

CC may be used to coat a medical device to reduce the S. aureus infection of
 CC an indwelling medical device, especially where the medical device is
 CC selected from the group consisting of vascular grafts, vascular stents,
 CC intravenous catheters, artificial heart valves, and cardiac assist
 CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or
 CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
 CC region or a gene encoding it may be used as an identifying probe for the
 CC identification of genes and encoding proteins from Staphylococcus aureus
 CC (other than Cifa), S. hemolyticus, S. lugdenensis, and S. schleriferi.
 CC The proteins of the invention have antibacterial activity.
 XX
 XX
 SQ Sequence 1166 AA;

Query Match 18.7%; Score 323.5; DB 20; Length 1166;
 Best Local Similarity 31.2%; Pred. NO. 1e-14;
 Matches 111; Conservative 53; Mismatches 137; Indels 55; Gaps 18;
 QY 2 VAADAPAGTITNLTNVT-----VGIDSGTTPVPHQAGVYKLYGFSVPNSAVKGDTF 56
 TQ 269 VQAPAAVANNV-NDLIKVTQTKVGGDKONVAAAHGDKDIEYDTEFTIDNKKKGDPM 327
 QY 57 KI-----TVPKELNLTSTAKVPPIMAGD---QVLANGVID-SDGNVITFTDYVNTK 107
 DB 328 TINDKNVPSDL-----TDKNPDIDTPSGEVIAGKTFDKATKQITFTDYVDKY 380
 QY 108 DDVATLTPAYIDPENY--KKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTI 165
 DB 381 EDIKSRLLSYIDKTKVNPETSLNLTATAGKETSQN--VTVDYQDPMVHGDSNIQISF 438
 QY 166 DQIDKNTNYQTIVNP--SGDNVIAPVL-----TGNLKPNTSDNALIDQNTSKV 216
 DB 439 TKLDEKQTIQQIYVNPPLKKSATNTKVDIAGSQVDDYGNKLGNGS-TIID-QNTEIKV 496
 QY 217 YKVDNAADLSSEYF-NPENFEDVTSNVIITFPNPNOYKVEFNTPD-DQITTPYIVVNG 274
 DB 497 YKVNDDQQLPQSNRIYDFSOYEDVTSQ-----FDNKKFSNNVATLDFGDINSAYIIKVS 552
 QY 275 HIDPNSKGLD-----ALRST-LYGYSNIIWRSMWDNEVAFNNGSGGDIKDP 323
 DB 553 KYTPTSDGELDIAOGTSMTTDKYGY-----YNYAGYSNFIVTSNDTGGDGTVRP 603

RESULT 22
 AAY08640
 ID AAY08640 standard; protein: 918 AA.
 XX
 AC AAY08640;
 09-AUG-1999 (first entry)
 DE S. aureus ClfB protein.
 XX
 KW Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD;
 KW SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
 KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
 KW extracellular matrix; vascular graft; vascular stent; vaccine;
 KW intravenous catheter; artificial heart valve; cardiac assist device;
 KW antibacterial.
 XX
 OS Staphylococcus aureus.
 XX
 PN W09927109-A2.
 XX
 PD 03-JUN-1999.
 XX
 PF 25-NOV-1998; 98WO-US25246.
 XX
 PR 31-AUG-1998; 98US-0098427.
 XX
 PR 26-NOV-1997; 97US-0066815.
 XX
 PA (EIDH/) EIDHIN D N.
 PA (FORF/) FORFAS T/A BIORESEARCH IRELAND.
 PA (FOST/) FOSTER T J.

PA (HOOK/) HOOK M A O.
 PA (INH-) INHIBITEX INC.
 PA (JOSE-) JOSEFSSON E.
 PA (PATT/) PATTI J M.
 PA (PERK/) PERKINS S E.
 XX
 PI Eidin DN, Foster TJ, Hook MAO, Josefsson E, Patti JM;
 PI Perkins SE;
 XX
 DR WPI: 1999-357844/30.
 DR N-PSDB: AAX77591.
 XX
 PT Staphylococcus aureus fibrinogen-binding proteins for treating
 PT septicemia, osteomyelitis, mastitis or endocarditis
 XX
 PS Claim 2; Fig 5; 143pp; English.
 XX
 CC This invention describes novel Staphylococcus aureus fibrinogen-binding
 CC proteins that bind both the alpha and beta fibrinogen chains. The
 CC proteins (and their encoding nucleic acids are ClfB, SdrC, SdrD and
 CC SdrE). Staphylococcus aureus is thought to utilize fibrinogen to adhere
 CC to medical devices, binding proteins that bind both the alpha and beta
 CC fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as
 CC competitive inhibitors to block this binding. Antibodies against ClfB,
 CC SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.
 CC The proteins of the invention can be used in a pharmaceutical composition
 CC for the treatment of Staphylococcus aureus infection e.g. septicemia,
 CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of
 CC S. aureus to the extracellular matrix. The proteins or their fragments
 CC may be used to coat a medical device to reduce the S. aureus infection of
 CC an indwelling medical device, especially where the medical device is
 CC selected from the group consisting of vascular grafts, vascular stents,
 CC intravenous catheters, artificial heart valves, and cardiac assist
 CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or
 CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
 CC region or a gene encoding it may be used as an identifying probe for the
 CC identification of genes and encoding proteins from Staphylococcus aureus
 CC (other than Cifa), S. hemolyticus, S. lugdenensis, and S. schleriferi.
 CC The proteins of the invention have antibacterial activity.
 XX
 SQ Sequence 918 AA;

Query Match 18.0%; Score 312.5; DB 20; Length 918;
 Best Local Similarity 24.9%; Pred. NO. 4.5e-14;
 Matches 92; Conservative 69; Mismatches 132; Indels 77; Gaps 15;
 QY 6 APAAGCTDITNLTNVTGIDSGTTPVPHQAGVYKLYGFSVPNSAVKGDTFKITYPKELN 65
 DB 213 ADAGTNNVDKVTASNFKLEK-TTFDPNOSGNTFMAANFTVTDKVKSGDYF----- 262
 QY 66 LNGVTSTAKVPPIMAGDOVLANGVID-----SDGNVI----- 97
 DB 263 -----TAKLP-----DSLTCNGVDVYNSNNTMPIADIKSTNGDVAKATYDILTKTYT 311
 QY 98 YTFDYVNTKDDVATLTPAYIDPENVKYKGNVTPLATIGSTTANKTVLVDYK-----Y 153
 DB 312 FVFYDYVNNKENINGQFSLPLFTDRAPKSKCTDYDANIADENFNKNTITYNYSPIAGI 371
 QY 154 GKFNLSIKGTIDQIDKTN--NTYQTIYVNP-----GDNVIAPVLTGNLKNPTOSNALI 207
 DB 372 DRPNANISSQIIGVDITASGONTYKQTFVFNPKORVLGNTWV--YIKGYQDKIEBSGKV 429
 QY 208 DQONTSIKVYKVDNAADLSSEYFVNP--ENFEDVTNSV--NITFPNPQYKVEFNTPDQ 263
 DB 430 SATDTKLRIFEVNDTSKLSDSYADPNDSNLKVEVDQFNRIYEHFNPVASKIFG----D 485
 QY 264 ITTPYIVVNGHIDPNSKGLALRSLTYGYN-----SNIWRSMSWDNEVAFNNGSGDG 319
 DB 486 ITKTYVVLVEGHYDNTGKN---LKTQVIOENVDPVTNRDYSIFGWNENNVYRGSGSADG 542
 QY 320 IDKPVVPEQP 329
 DB 543 -DSAVNPKDP 551

QY 111 KATLTPAYIDPENKKTGNVTLATIGTSTANKTVLVDY-----EKYKFNLSIKGPI 165
 Db 132 TAEISLNLFDPTTQKGNQNVKLGITTTYSKIFNIOYLGVRDNG-----VTANGRI 187
 QY 166 DOIDKNTNTYROTIVNPSGDNVIAPIVLTGNL-----KPNVDSNALIDQONTSIKYKYVDN 221
 Db 188 DTLNKRVDGKFSHFAYMKPNQSLSSVTGTQVTKGNKPGVN-----NPTVKVYKHIG 239
 QY 222 RADLSES 228
 Db 240 SDDLAE 246
 RESULT 24
 AAY08642
 ID AAY08642 standard; Protein; 1315 AA.
 XX AC
 XX XX
 DT 09-AUG-1999 (first entry)
 XX XX
 DE S. aureus SdrD protein.
 XX KW Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD;
 KW SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
 KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
 KW extracellular matrix; vascular graft; vascular stent; vaccine;
 KW intravenous catheter; artificial heart valve; cardiac assist device;
 KW antibacterial.
 XX OS Staphylococcus aureus.
 XX PN WO9927109-A2.
 XX PD 03-JUN-1999.
 XX PF 25-NOV-1998; 98WO-US25246.
 XX PR 31-AUG-1998; 98US-0098427.
 XX PR 26-NOV-1997; 97US-0068815.
 XX (EIDH/) EIDHIN D N.
 PA (FORF-) FORFAS T/A BIORESEARCH IRELAND.
 PA (FOST/) FOSTER T J.
 PA (HOOK/) HOOK M A O.
 PA (INH-) INHIBITEX INC.
 PA (JOSE/) JOSEFSSON E.
 PA (PATT/) PATTI J M.
 PA (PERK/) PERKINS S E.
 XX Eidhin DN, Foster TJ, Hook MAO, Josefsson E, Patti JM;
 PI Perkins SE;
 XX WPI; 1999-357844/30.
 DR N-PSDB; AAX77593.
 XX Staphylococcus aureus fibrinogen-binding proteins for treating
 PT septicemia, osteomyelitis, mastitis or endocarditis
 XX Claim 8; Fig 8; 143pp; English.
 XX This invention describes novel Staphylococcus aureus fibrinogen-binding
 CC proteins that bind both the alpha and beta fibrinogen chains. The
 CC proteins (and their encoding nucleic acids) are ClfB, SdrC, SdrD and
 CC SdrE. Staphylococcus aureus is thought to utilize fibrinogen to adhere
 CC to medical devices, binding proteins that bind both the alpha and beta
 CC fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as
 CC competitive inhibitors to block this binding. Antibodies against ClfB,
 CC SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.
 CC The proteins of the invention can be used in a pharmaceutical composition
 CC for the treatment of Staphylococcus aureus infection e.g. septicemia,
 CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of
 CC S. aureus to the extracellular matrix. The proteins or their fragments

AAW89804
 ID AAW89804 standard; Protein; 251 AA.
 XX AC
 XX AAW89804;
 DT 16-MAR-1999 (first entry)
 XX XX
 DE Staphylococcus aureus protein SEQ ID #5252.
 XX KW Computer readable medium; vaccine; S. aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome.
 XX OS Staphylococcus aureus.
 XX EP786519-A2.
 DT 30-JUL-1997.
 XX PF 07-JAN-1997; 97EP-0100117.
 XX PR 05-JAN-1996; 96US-0009861.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 PI Rosen CA;
 XX WPI; 1997-374922/35.
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus
 PT stored on computer readable medium and used in the production of
 PT anti-S. aureus vaccines
 XX Claim 23; Page 3261-3262; 3271pp; English.
 XX This sequence represents a Staphylococcus aureus protein sequence of the
 CC invention. The DNA sequences encoding the S. aureus proteins are recorded
 CC on a computer readable medium, preferably selected from a floppy or hard
 CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.
 CC Homology searches using the S. aureus DNA sequences allows putative
 CC functions to be assigned so that protein-encoding or regulatory regions
 CC of commercial, therapeutic or industrial importance can be obtained.
 CC Specifically, sequences which are likely to encode antigens have been
 CC identified and these polypeptides can be used in a vaccine composition
 CC against S. aureus infection. The polypeptides can also be used in a kit
 CC for the immunodetection of S. aureus in a sample. S. aureus is implicated
 CC in numerous human diseases, including cellulitis, eyelid infections, food
 CC poisoning, osteomyelitis, skin and surgical wound infections, scalded
 CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the
 CC DNA sequences can be used for recombinant production of the polypeptides.
 CC The new DNA sequences (and their fragments) are useful as primers or
 CC probes for isolating homologues of any of the 5191 S. aureus DNA sequences
 CC contained on the computer readable medium.
 XX SQ Sequence 251 AA;
 Query Match 16.3%; Score 282.5; DB 18; Length 251;
 Best Local Similarity 29.1%; Pred. No. 1.2e-12;
 Matches 72; Conservative 46; Mismatches 92; Indels 37; Gaps 7;
 QY 2 VAADAPAGTDTITNLTNTVTVGIDSGT-----TYPHQAGYKLVKNGFSPVNSAV 51
 Db 17 VVEETKATGTDVTKN-----VEVEGSEIVGHKQDTNVNPNRAERVTLYKKWKEGK 71
 QY 52 KGDTFTKIVPRLNGLNGVTSTAKVPPINAGD-QVLANGVIDSDGNVIYTFDYNVTKDDV 110
 Db 72 AGDYFDFTLSNVETHGISTLRKVPKSTQGVNATGEIIGERKVRVTFKRYVOEKD 131

CC may be used to coat a medical device to reduce the *S. aureus* infection of
CC an indwelling medical device, especially where the medical device is
CC selected from the group consisting of vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves, and cardiac assist
CC devices. ClfB, SdrC, SdrE, or an active fragment, subdomain or
CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
CC region of a gene encoding it may be used as an identifying probe for the
CC identification of genes and encoding proteins from *Staphylococcus aureus*
CC (other than CrlA), *S. hemolyticus*, *S. lugdenensis*, and *S. schleiferi*.
CC The proteins of the invention have antibacterial activity.
XX
XX

SQ Sequence 1315 AA;

Query Match 14.9%; Score 257.5; DB 20; Length 1315;

Best Local Similarity 26.9%; Pred. No. 6.5e-10;

Matches 88; Conservative 51; Mismatches 157; Indels 31; Gaps 12;

QY 13 ITNQLTNTVGTDSGTTVYPHQAGYVKNLYGFSVPNSAVKGDFTKIVPKELNLNGVT-- 70

251 ITSNTTLTVVDADKNNKIVPAQ-DYLSLKSQITVDDKVKSGDYFTIKYSDTVQVYGLNPE 309

71 -----STAKVPPIMAGDQVLANGVIDSDGNVI-YTFTDYVNTKDDVKATLTMPAYIDPEN 124

310 DIKNIGDIKDP---NNGETIATAKHDTANNLITYTFTDYVDRFNSVQMGINSIYADAT 366

QY 125 VKKTGN-VTLATGIGSTTANKTVLDYKGYKFNLSIKG----TIDQIDKTNNT--YRQ 177

367 IPVSKNDVEFNVTIGNTTKTANIQYDPDYVYVNEKNSIGSAFTETVSHVGNKENPGYKQ 426

QY 178 TIYVNPNGDNVIAPVLITGNLKPNTDSNAL--IDQONTSIKVKVDNAADLSSEYFVNPEN 235

427 TIYVNPSENLSNAKLKVOAYHSSYPNNIGQINKDVTDIKIQVPGYTLNKGVDYNTKE 486

QY 236 FEDVTNSV--NITFPNPQYKVEFNTPDDQITPYIVVNGHID-PNSKGDALALRSTLYG 292

487 LTDVTNQYLQKITYGDNNSAVIDFGNAD-----SAYVVMYNTKFOYTNSSEPTLVQMATLS 542

QY 293 YNSNIWRSMWSNDNEVAFNNGSGGDG 319

543 STGN---KSVSTGNALGFTNNOSGGAG 566

RESULT 25

AAU34402

ID AAU34402 standard; Protein; 1349 AA.

AC AAU34402;

14-FEB-2002 (first entry)

Staphylococcus aureus cellular proliferation protein #678.

Antisense: prokaryotic cellular proliferation protein;

antibiotic; antibacterial; drug design.

Staphylococcus aureus.

WO200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US09180.

21-MAR-2000; 2000US-191078P.

23-MAY-2000; 2000US-206848P.

26-MAY-2000; 2000US-207727P.

23-OCT-2000; 2000US-242578P.

27-NOV-2000; 2000US-253625P.

22-DEC-2000; 2000US-257931P.

16-FEB-2001; 2001US-269308P.

(ELIT-) ELITRA PHARM INC.

XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR N-PSDB; AAS52261.
XX

PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX

PS Example 3; Seq ID No 5898; 51lpp; English.

XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, themselves in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1349 AA;

Query Match 14.9%; Score 257.5; DB 22; Length 1349;

Best Local Similarity 26.9%; Pred. No. 6.7e-10;

Matches 88; Conservative 51; Mismatches 157; Indels 31; Gaps 12;

QY 13 ITNQLTNTVGTDSGTTVYPHQAGYVKNLYGFSVPNSAVKGDFTKIVPKELNLNGVT-- 70

251 ITSNTTLTVVDADKNNKIVPAQ-DYLSLKSQITVDDKVKSGDYFTIKYSDTVQVYGLNPE 309

QY 71 -----STAKVPPIMAGDQVLANGVIDSDGNVI-YTFTDYVNTKDDVKATLTMPAYIDPEN 124

310 DIKNIGDIKDP---NNGETIATAKHDTANNLITYTFTDYVDRFNSVQMGINSIYADAT 366

QY 125 VKKTGN-VTLATGIGSTTANKTVLDYKGYKFNLSIKG----TIDQIDKTNNT--YRQ 177

367 IPVSKNDVEFNVTIGNTTKTANIQYDPDYVYVNEKNSIGSAFTETVSHVGNKENPGYKQ 426

QY 178 TIYVNPNGDNVIAPVLITGNLKPNTDSNAL--IDQONTSIKVKVDNAADLSSEYFVNPEN 235

427 TIYVNPSENLSNAKLKVOAYHSSYPNNIGQINKDVTDIKIQVPGYTLNKGVDYNTKE 486

QY 236 FEDVTNSV--NITFPNPQYKVEFNTPDDQITPYIVVNGHID-PNSKGDALALRSTLYG 292

487 LTDVTNQYLQKITYGDNNSAVIDFGNAD-----SAYVVMYNTKFOYTNSSEPTLVQMATLS 542

QY 293 YNSNIWRSMWSNDNEVAFNNGSGGDG 319

543 STGN---KSVSTGNALGFTNNOSGGAG 566

Search completed: June 23, 2003, 14:01:38

Job time : 56.406 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:00:06 ; Search time 19.367 Seconds
(without alignments)

502.864 Million cell updates/sec

Title: US-10-056-052A-4

Perfect score: 1732

Sequence: 1 MVAADAPAAAGTDITNLTNV.....NGSGGDGIDKPVVPEQDE 331

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

1 number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCITUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1732	100.0	345	4	US-08-856-253-7
2	1727	99.7	933	3	US-08-293-728-2
3	1727	99.7	933	4	US-09-421-868-2
4	353	20.4	930	4	US-09-134-001C-5314
5	204	11.8	1112	2	US-08-714-402-2
6	204	11.8	1161	4	US-09-327-536-2
7	121.5	7.0	992	4	US-09-206-942-61
8	121.5	7.0	998	4	US-09-206-942-59
9	120	6.9	737	4	US-09-071-035-460
10	120	6.9	2032	4	US-09-071-035-458
11	120	6.9	2032	4	US-09-071-035-462
12	120	6.9	2032	4	US-09-071-035-466
13	118	6.8	789	4	US-09-002-285-94
14	117	6.8	1180	4	US-09-206-942-65
15	117	6.8	1188	4	US-09-206-942-63
16	116	6.7	789	4	US-09-002-285-96
17	115	6.6	790	4	US-08-960-780-4
18	115	6.6	790	4	US-09-073-898-4
19	114	6.6	789	1	US-08-471-033-29
20	114	6.6	789	2	US-08-471-044-29
21	114	6.6	789	2	US-08-463-483A-29
22	114	6.6	789	2	US-08-471-046A-29
23	114	6.6	789	2	US-08-470-566B-29
24	114	6.6	789	2	US-08-838-219B-2
25	114	6.6	789	2	US-08-469-334-29
26	114	6.6	789	3	US-09-300-529-29
27	114	6.6	789	3	US-09-233-336A-2

28	114	6.6	789	4	US-09-233-752A-2
29	114	6.6	789	4	US-09-402-036-2
30	114	6.6	789	4	US-09-002-285-78
31	114	6.6	789	4	US-09-002-285-100
32	114	6.6	789	4	US-09-904-226-2
33	114	6.6	790	4	US-08-960-780-8
34	114	6.6	790	4	US-09-073-898-8
35	113	6.5	789	4	US-08-960-780-6
36	113	6.5	789	4	US-09-073-898-6
37	113	6.5	789	4	US-09-002-285-80
38	113	6.5	789	4	US-09-002-285-90
39	113	6.5	1073	4	US-09-206-942-49
40	113	6.5	1079	4	US-09-206-942-47
41	113	6.5	1095	4	US-09-206-942-45
42	113	6.5	1101	4	US-09-206-942-43
43	113	6.5	1394	4	US-08-296-791-2
44	113	6.5	1394	5	PCT-US95-10661A-2
45	112	6.5	759	4	US-09-002-285-86
46	112	6.5	789	4	US-09-002-285-84
47	111.5	6.4	669	4	US-09-071-035-264
48	111.5	6.4	1004	4	US-09-206-942-57
49	111.5	6.4	1010	4	US-09-206-942-55
50	111.5	6.4	1638	4	US-09-071-035-258
51	111.5	6.4	1638	4	US-09-071-035-262
52	111.5	6.4	1638	4	US-08-838-219B-6
53	111	6.4	746	3	US-09-233-752A-6
54	111	6.4	746	4	US-09-402-036-6
55	111	6.4	746	4	US-09-904-226-6
56	111	6.4	746	4	US-08-447-031A-2
57	111	6.4	746	4	US-08-728-470-9
58	110	6.4	1338	2	US-08-719-641-9
59	110	6.4	1338	4	US-08-617-697-9
60	110	6.4	1599	2	US-09-206-942-41
61	110	6.4	1599	2	US-09-206-942-39
62	109.5	6.3	1005	4	US-09-002-285-88
63	109.5	6.3	1011	4	US-08-471-033-32
64	109	6.3	511	4	US-08-471-033-32
65	109	6.3	789	1	US-08-463-483A-32
66	109	6.3	789	2	US-08-471-046A-32
67	109	6.3	789	2	US-08-470-566B-32
68	109	6.3	789	2	US-08-838-219B-4
69	109	6.3	789	2	US-08-469-334-32
70	109	6.3	789	3	US-09-300-529-32
71	109	6.3	789	3	US-09-233-336A-4
72	109	6.3	789	4	US-09-402-036-4
73	109	6.3	789	4	US-09-002-285-82
74	109	6.3	789	4	US-09-002-285-92
75	109	6.3	789	4	US-09-002-285-98
76	109	6.3	789	4	US-09-904-226-4
77	109	6.3	789	4	US-09-453-702B-257
78	109	6.3	789	4	US-09-002-285-102
79	109	6.3	789	4	US-08-169-927-2
80	108.5	6.3	3169	4	US-08-621-944A-4
81	108	6.2	790	4	US-08-945-567D-4
82	108	6.2	1612	1	US-08-621-944A-3
83	108	6.2	1833	4	US-08-945-567D-3
84	108	6.2	1833	4	US-09-268-347-48
85	108	6.2	1992	4	US-09-206-942-35
86	108	6.2	1992	4	US-09-206-942-37
87	108	6.2	2048	4	US-09-206-942-34
88	107.5	6.2	915	4	US-08-856-253-6
89	107.5	6.2	1228	4	US-08-894-017-23
90	107.5	6.2	1228	4	US-08-184-748-2
91	107	6.2	512	3	US-08-414-926A-5
92	106	6.1	1561	3	US-08-926-922-5
93	105	6.1	721	4	US-09-253-682-5
94	105	6.1	1565	6	US-09-527-657-5
95	104.5	6.0	399	1	US-08-266-311-1
96	104.5	6.0	399	2	US-08-467-527A-1
97	104.5	6.0	399	3	
98	104.5	6.0	399	4	
99	104.5	6.0	870	2	
100	104.5	6.0	870	2	

Patent No. 5352450

ALIGNMENTS

RESULT 1
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

US-08-856-253-7

Query Match 100.0%; Score 1732; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-143;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAADAPAAAGTDITNLTNTVTVGIDSGTIVYVPHQAGYVKNLYGFSVPNSAVKGDFTKITV 60
DB 13 MVAADAPAAAGTDITNLTNTVTVGIDSGTIVYVPHQAGYVKNLYGFSVPNSAVKGDFTKITV 72
QY 61 PRELNLTGVTSTAKVPPIMAGDOVLANGVIDSDGNVYFTFDVNTKDDVKATLTMPAYI 120
DB 73 PRELNLTGVTSTAKVPPIMAGDOVLANGVIDSDGNVYFTFDVNTKDDVKATLTMPAYI 132
QY 121 DPENVKKTGNVLTATGISTTANKTANLVLDYKYGKFNLSIKGTIDQIDKTNNTYRQTIY 180
DB 133 DPENVKKTGNVLTATGISTTANKTANLVLDYKYGKFNLSIKGTIDQIDKTNNTYRQTIY 192
QY 181 VNPSSGDNVLTAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVT 240
DB 193 VNPSSGDNVLTAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVT 252
QY 241 NSVNITFPNPNOYKVEFNTPDDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYSNLIWR 300

DB 253 NSVNITFPNPNOYKVEFNTPDDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYSNLIWR 312
QY 301 SMSWDNEVAFNNGSGSGDIDKPVVPEQDPE 331
DB 313 SMSWDNEVAFNNGSGSGDIDKPVVPEQDPE 343

RESULT 2

US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match 99.7%; Score 1727; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 7.1e-142;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAADAPAAAGTDITNLTNTVTVGIDSGTIVYVPHQAGYVKNLYGFSVPNSAVKGDFTKITV 61
DB 221 VAADAPAAAGTDITNLTNTVTVGIDSGTIVYVPHQAGYVKNLYGFSVPNSAVKGDFTKITV 280
QY 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVYFTFDVNTKDDVKATLTMPAYID 121
DB 281 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVYFTFDVNTKDDVKATLTMPAYID 340
QY 122 PENVKKTGNVLTATGISTTANKTANLVLDYKYGKFNLSIKGTIDQIDKTNNTYRQTIY 181
DB 341 PENVKKTGNVLTATGISTTANKTANLVLDYKYGKFNLSIKGTIDQIDKTNNTYRQTIY 400
QY 182 NPSSGDNVLTAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 241
DB 401 NPSSGDNVLTAPVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 460
QY 242 SVNITFPNPNOYKVEFNTPDDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYSNLIWR 301
DB 461 SVNITFPNPNOYKVEFNTPDDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYSNLIWR 520
QY 302 MSWDNEVAFNNGSGSGDIDKPVVPEQDPE 331
DB 521 MSWDNEVAFNNGSGSGDIDKPVVPEQDPE 550

RESULT 3

US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421,868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933


```

Db 458 RYSAKETNWNISNGDREGSTIIDSTTIKVKYKGDNONLDPNSRIYDYSEVEDVTNDYA 517
Qy 246 TFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGD-----LALRSTLYGYSNII 298
Db 518 QLGNNNDVINFG---NIDSPYIKIVISKYDPN-KDDYTTIQOVTVMQTTINEYGE-- 570
Qy 299 WRMSWDNEVAFNNGSGSGDGIDKPVVPEQ 328
Db 571 FRTASYDNTIAFTSSGGQGG-DLP--PEK 597

RESULT 5
US-08-714-402-2
; Sequence 2, Application US/08714402
; Patent No. 5910441
; GENERAL INFORMATION:
; APPLICANT: ROCHA, Claudia
; APPLICANT: FISCHETTI Vincent A.
; TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING
; TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,402
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 016921-097
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1112 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-714-402-2

Query Match 11.8%; Score 204; DB 2; Length 1112;
Best Local Similarity 22.8%; Pred. No. 3.3e-09;
Matches 86; Conservative 72; Mismatches 145; Indels 74; Gaps 18;

Qy 8 AAGCTDITN--OLTNVTVGIDS--CCT-----VYPHQAGYVVLNKGFCFSPNSAV---K 52
Db 142 AGSKDVSSSLQLENPKRMSVSKYKTEVSSGAADFYRNHAAFYKMSPELKQKDKSETNP 201
Qy 53 GDTFKIVPKELNGLVNTSTAKVPIM--AGDQVLANGVIDSCGN-VIYTFDVTNNTKDD 109
Db 202 GDTFVLQDLRRLNPKGLISQ--DIPKLIYDSANSPLAIGKYHAENHQLIYTFDVIAGLDK 259
Qy 110 VKATLTNPAVIDPE-----NVKKT-----GNVTLATGIGSTTANKTVLDVYE 151
Db 260 VLSAELSLEFLENKEVLENTSISNFKSTIGGQETTKYKGTNVLYGNSTKESNYITNGLS 319
Qy 152 KYGKFYNLSTKGTIDQIDKTNNTVROLYIYNPSSGDNVIAPVLITGNL-----KPNTDSNALI 207
Db 320 NVG-----GSIESYNTETGFEFVYVYVNPNTNI--PYATNMLMGFGARSNTSDLE 369

```


QY 19 NVTG-GIDSGTTPVPHQAGYKLVNFGSPVNSAVKGDTEKTVPKELNGLNGVTAKVPP 77
 DB 405 NVTGQDSSSTI-----TGNINISQAANVTURAYNGNCRN-----KQLTGLNVSIEGNSL 456
 QY 78 IMAGDOVLN-----GVIDSDGNVIYTFDDYNTKDDVKATLTPAYIDPENVK 127
 DB 457 IGASANINGLSVENAKFKGETOONLITGTFINNGDSKNIS-----QGVVK 505
 QY 128 TGNVT-----LATG-----IGSTTANKTVLVDEYKGYKFNLSIKGTIDQ----- 167
 DB 506 LGNVTNDGLNITTHAKHNORSIIGGDIINKGSLNITDSKNKAEIQIGGNSIKNGLT 565
 QY 168 --IDKNTNVTQTYVNPSPGDNVIAPVLTGNLKPNTDSNALIDQNTSIKYKYVDNAADL 225
 DB 566 ISSDKINITNQTITKAGVDGE-----NSDSDA-TNNANLTIKTKELKLTODL 611
 QY 226 SESYFVNPE-----NPEVDVNSVNTFPNPQYKVEFFTPDQIIT-PIV 270
 DB 612 NISGFNKAETAKGSDLTIGNTNSADSTNAKKVTF-----NOVK-----DSKISAGDHN 662
 QY 271 VVNHGIDPNKGDALRSTLYGNSNIIRWSMWDNEVAFNNGSGSD 318
 DB 663 TLNSKVETSGNTD-----NTGDGSGN 683

RESULT 9
 US-09-071-035-460
 ; Sequence 460, Application US/09071035
 ; Patent No. 6448043
 ; GENERAL INFORMATION:
 ; APPLICANT: GIL H. Choi
 ; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
 ; NUMBER OF SEQUENCES: 496
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/071.035
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: A. Anders Brookes
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB369P2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 460:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 737 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-071-035-460

Query Match 6.9%; Score 120; DB 4; Length 737;
 Best Local Similarity 21.9%; Pred No. 0.037;
 Matches 79; Conservative 54; Mismatches 139; Indels 88; Gaps 18;

QY 8 AAGTDITNQLT-NVTVGIDS-----GTTVYPHQAGYKVL--NYGFSVPNSAVK-- 52
 DB 172 AGPTDISDYFGDETTIIDNEDPIYLNPDCTPATPYKEDVTIHWNFNWSIPEDVRQOM 231
 QY 53 --GDTFKITVPKELNGLNGVTSTAKVPPIMAGDOVLN-----VIDSDGNVIYTFDDY 104
 DB 232 KAGDYFEFQLPGNKLKPNKPGS-----GDLVDAEGNVYGTVTISDGTGRTFTFNRI 282
 QY 105 NTKDDVKATLTPMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDEYKGYKFNLSIKGT 164
 DB 283 TSESDIHCDFSLDTHLNDSDGRGPDWVI--DIPTQEDLPPVVIPI-----VPDT 330
 QY 165 IDQIDKNTNVTQTYVNPSPGDNVIAPVLTGNLKPNTDSNALIDQNTSIKYKYVDNAADL 211
 DB 331 EQIDKQCHGDFRTP--NPS-----AITWTVDIN--QAMKDOTNPVTETWPTGNTF 377
 QY 212 TSIKYKYVDNAAD--LSE-SYFVNPE--DVTNSVNTFPNPQYKVEFFTPDQIITP 267
 DB 378 KSVKYELVLMNLDGTIREVGRLESPDEYTVOKNGNVTIKGTNKAIRLEYOTTIDEVIP 437
 QY 268 ---YIVVYNGHI---DPNSKGDALRSTLYGNSNIIRWSMWDN-----EVAFNNG 313
 DB 438 DGGGDVPEPKHATLTSDNNPNGLDAEATVTATYCKMLDKRNDYDEANQEFWEINYNYG 497

RESULT 10
 US-09-071-035-458
 ; Sequence 458, Application US/09071035
 ; Patent No. 6448043
 ; GENERAL INFORMATION:
 ; APPLICANT: GIL H. Choi
 ; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
 ; NUMBER OF SEQUENCES: 496
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/071.035
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: A. Anders Brookes
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB369P2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 458:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2032 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-071-035-458

Query Match 6.9%; Score 120; DB 4; Length 2032;
 Best Local Similarity 21.9%; Pred No. 0.17;
 Matches 79; Conservative 54; Mismatches 139; Indels 88; Gaps 18;
 QY 8 AAGTDITNQLT-NVTVGIDS-----GTTVYPHQAGYKVL--NYGFSVPNSAVK-- 52

Db 204 AGPTDISDYFTGDETTIIDNFEDPIYLNPDGTPATPPYKEDVTIHWNFNWSIPEDVREOM 263
QY 53 --GDTFKITVPKELNGLNGVTSTAKVPPIMAGDOVLANG-----VIDSGNVYITFTDYV 104
Db 264 KAGDYFEFQLPGLNPKNPGS-----GDLVDAEGNVYGTYTISEDGTVRTFTNERI 314
QY 105 NTKDDVKATLMPAYIDPENVKKTGNVTLAGTIGSTTANKTVLVDYKYGKFNLSIKGT 164
Db 315 TSEDHIGDFSLDHLNDSGRGPDGWI--DIPQEDLPPVPIPI-----VPDT 362
QY 165 IDQDKTNTTYRTIYVNPSSGDNVAPVLTGNLKPNTDSNALIDQON----- 211
Db 363 EQQIDKQGHEDRTP--NPS-----AITWTVDIN--QAMKQDNTVTVETWPTGNTF 409
QY 212 TSIVKVKVDNAAD--LSE-SYFVNPFENFE-DVTSVNTFTFPNPNQYKVEFNTPDDQIITP 267
Db 410 KSVKYVELVMLNDGTIKEVGRLESPDEYTVDKNGNVTIKGDTNKAYRLEYQTTIDEAVIP 469
268 ---YIVVNGHI-----DPNSKGDALRSTLYGYSNIIWRSMWDN-----EVAFNNG 313
470 DGGGDVPEKNHATLTSDNPNGLDAEATVATYCKMLDKRNLIDYEANQEFWEINYNYG 529

RESULT 11

US-09-071-035-462
; Sequence 462, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: GIL H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 462:
SEQUENCE CHARACTERISTICS:
LENGTH: 2032 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-462

Query Match 6.9%; Score 120; DB 4; Length 2032;

Best Local Similarity 21.9%; Pred. No. 0,17;

Matches 79; Conservative 54; Mismatches 139; Indels 88; Gaps 18;
QY 8 AAGDTITNQLT-NVTVGIDS-----GTTVYPHOAGYVKL--NYGFSVPNSAVK-- 52
Db 204 AGPTDISDYFTGDETTIIDNFEDPIYLNPDGTPATPPYKEDVTIHWNFNWSIPEDVREOM 263

Db 204 AGPTDISDYFTGDETTIIDNFEDPIYLNPDGTPATPPYKEDVTIHWNFNWSIPEDVREOM 263
QY 53 --GDTFKITVPKELNGLNGVTSTAKVPPIMAGDOVLANG-----VIDSGNVYITFTDYV 104
Db 264 KAGDYFEFQLPGLNPKNPGS-----GDLVDAEGNVYGTYTISEDGTVRTFTNERI 314
QY 105 NTKDDVKATLMPAYIDPENVKKTGNVTLAGTIGSTTANKTVLVDYKYGKFNLSIKGT 164
Db 315 TSEDHIGDFSLDHLNDSGRGPDGWI--DIPQEDLPPVPIPI-----VPDT 362
QY 165 IDQDKTNTTYRTIYVNPSSGDNVAPVLTGNLKPNTDSNALIDQON----- 211
Db 363 EQQIDKQGHEDRTP--NPS-----AITWTVDIN--QAMKQDNTVTVETWPTGNTF 409
QY 212 TSIVKVKVDNAAD--LSE-SYFVNPFENFE-DVTSVNTFTFPNPNQYKVEFNTPDDQIITP 267
Db 410 KSVKYVELVMLNDGTIKEVGRLESPDEYTVDKNGNVTIKGDTNKAYRLEYQTTIDEAVIP 469
268 ---YIVVNGHI-----DPNSKGDALRSTLYGYSNIIWRSMWDN-----EVAFNNG 313
470 DGGGDVPEKNHATLTSDNPNGLDAEATVATYCKMLDKRNLIDYEANQEFWEINYNYG 529

RESULT 12

US-09-071-035-466
; Sequence 466, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: GIL H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 466:
SEQUENCE CHARACTERISTICS:
LENGTH: 2032 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-466

Query Match 6.9%; Score 120; DB 4; Length 2032;

Best Local Similarity 21.9%; Pred. No. 0,17;

Matches 79; Conservative 54; Mismatches 139; Indels 88; Gaps 18;
QY 8 AAGDTITNQLT-NVTVGIDS-----GTTVYPHOAGYVKL--NYGFSVPNSAVK-- 52
Db 204 AGPTDISDYFTGDETTIIDNFEDPIYLNPDGTPATPPYKEDVTIHWNFNWSIPEDVREOM 263

QY 53 --GDTFKITVPKELNLTSTAKVPPIMAGDOVLANG-----VIDSDGNVIYFTDYV 104
 Db 264 KAGDYFEQPCNLPKPGS-----GDLVDAEGNYGYTYISEDGTVRTFNERI 314
 QY 105 NTKDDVKATLMPAIDENVKKTGNVTLATGIGSTTANKTVLVLDYKGYFNLSIGKT 164
 Db 315 TSEIDHGFSLDTHLNSDGRGPDGWI--DIPTQEDLPPVVIPI-----VPDT 362
 QY 165 IDIQDKTNTTYRTIYVNPSPGDNVAPVLTGKLNKPTDSNALIDQON-----211
 Db 363 EQQIDKQGHFDRTP---NPS-----AITWTVDIN---QAMKQDTNVTVTWTGTGTF 409
 QY 212 TSIKYKVDNAD--LSE-SYFVNPFENFE-DVTSNVNITFPNPNQYKVEFNTPDDQITTP 267
 Db 410 KSVKVELVNLMDGTIKEVRELSPEYTVOKNGHVTIKGDTNKAYRLEYQTTIDEAVIP 469
 QY 268 ---YIVVNGHI---DPSKGDALRLSTLYGYSNIIWRSMWDN-----EYAFNNG 313
 470 DGGGVPPKHNATLTSNPNGLDAEATVTATYGMKLDKRNIDYDEANOEFTWEINNYG 529

RESULT 13
 US-09-002-285-94
 ; Sequence 94, Application US/09002285
 ; Patent No. 6369213

GENERAL INFORMATION:
 ; APPLICANT: Schnepf, H. Ernest
 ; APPLICANT: Wicker, Carol
 ; APPLICANT: Narva, Kenneth E.
 ; APPLICANT: Walz, Michelle
 ; APPLICANT: Stockhoff, Brian
 ; APPLICANT: Muller-Cohn, Judy
 ; TITLE OF INVENTION: Toxins Active Against Pests
 ; NUMBER OF SEQUENCES: 105
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: Florida
 ; COUNTRY: USA
 ; ZIP: 32606

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/002,285
 ; FILING DATE:

CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/886,615
 ; FILING DATE: 1-JUL-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/674,002
 ; FILING DATE: 1-JUL-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sanders, Jay M.
 ; REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-701C2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (352) 375-8100
 ; TELEFAX: (352) 372-5800
 ; INFORMATION FOR SEQ ID NO: 94:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 789 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-002-285-94

Query Match 6.8%; Score 118; DB 4; Length 789;
 Best Local Similarity 22.0%; Pred. No. 0.061;
 Matches 82; Conservative 56; Mismatches 133; Indels 102; Gaps 18;
 QY 12 DITNOLTNVTVGIDSGTTVYPHQAGYVKNLYGFSVPNSAVKGDTPKITYPKELNLTNGVTS 71
 Db 148 EISDKLDIINNVNLINSLTLEITPAYQRIKY-----VNEKFEELTFATETSSKVKKDG--S 201
 QY 72 TAKVPPIMAGDOVLANGVIDSGNVIYFTDYVNTKDDV-----KATLTMPAYIDP 122
 Db 202 PADILDELAELTELAKSVTKND---VDGFEFLNTFHDVWVGNLFGRSALKTASELITK 258
 QY 133 ENVKKTGN-----VTLATIGS-----TTANKTV---LVDY-----EKYKPY 157
 Db 259 ENVKTSGEVGNVNFLLVLTALQAKAFLTLTCRKLGLADIDVTSIMNEHLNKEKEEF 318
 QY 158 NLSIKGTIDIDKTNNTYRTIYVNPSPGDNVAPVLTGKLNKPTDSNALI--DOQNTSJK 215
 Db 319 RVNHLPTL-----SNTFSNPNYAKVKGSDDEAKMIV-EAKP---GHALIGFEISNDSIT 368
 QY 216 VYKVDNAADLSSEYFVN-----PENFEDVTSNVNITFPNPNQY---KV 255
 Db 369 VLKV-YEAKLKONYQVDKDSLEVIYGDMDKLLCPDQSEQIYTTNNIVF--PNEIVITKI 425
 QY 256 EF-----NTPDDQITTPYIVVNGHIDPNKSGDLALRLSTLYGYSNIIWRSMWDNEVAFN 311
 Db 426 DFTKKMKTLYEVTANFVDSSTGEIDLNNK-----KVESSEAEYR 465
 QY 312 NCGSGDGDIDKRPV 324
 Db 466 TLSANDDGVYMLP 478

RESULT 14
 US-09-206-942-65
 ; Sequence 65, Application US/09206942
 ; Patent No. 6432669

GENERAL INFORMATION:
 ; APPLICANT: Loosmore, Sheena M.
 ; APPLICANT: Yang, Yan-Ping
 ; APPLICANT: Klein, Michel H.
 ; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
 ; FILE OF INVENTION: Molecular Weight Proteins
 ; FILE REFERENCE: 1038-861 MIS:jb
 ; CURRENT APPLICATION NUMBER: US/09/206,942
 ; CURRENT FILING DATE: 1998-12-08
 ; EARLIER APPLICATION NUMBER: 09/167,568
 ; EARLIER FILING DATE: 1998-10-07
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 65
 ; LENGTH: 1180
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-09-206-942-65

Query Match 6.8%; Score 117; DB 4; Length 1180;
 Best Local Similarity 22.3%; Pred. No. 0.14;
 Matches 87; Conservative 49; Mismatches 143; Indels 112; Gaps 20;
 QY 3 AADAPAACTDITNLTNTVTVG--IDSGTTVYPHQAGYVKNLYGFSVPNSAVKGDTPKITY 60
 Db 29 SADTPKRNNTKTTLTNSTLEKILARGSPVNTANNEIRVNSDINI-----GGNSHLTL 82
 QY 61 PKELNLTNGVTSTAKVPPIMAGDOVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYI 120
 Db 83 WSSKNK-----SGVLINGNITSTAN--GNLTYSYSGWVDIHKNTILES-- 124
 QY 121 DPENV-KKTGNVTLATIGST--TANKTVLVLDYKGYFNLSIGKT-----IDQIDKT 171
 Db 125 GRLNITTKEGDVAFKGNLNTITGGOTTITAGNNKGFRENVSLNGTGTGLLENLRPKOKN 184
 QY 172 N-----NTYRTIYVNPSPGD-NVIAPVLGTN-----LKPNTDSNALIDQON 211

Db 185 NSLVTNYEGLNLSGVSNIEMIPNATSNWYSRYKGRTYWNITHLNASDSNF-----N 239
QY 212 TSIIKVKVDNADLSSEYFVNPENEDVTN-----SVNITFPNP-----NOYK----- 254
Db 240 LTIDSSAEDGSAPLLSSYTLNGISFTTDTTFNVNNAKVNFIKAPICTINOYNLNLVAL 299
QY 255 -----VEFTPDQITTPYVVVNGHIDPNSKGLALR-----STLYGYNS 295
Db 300 FNGNLSVSGGNNVTRLNASSNQOTPGVIINSKHLNA-SKGS-SLRETTGCTKVGFLI 357
QY 296 NIIWRMSWDNEAFNNGSGS-----GDGID 321
Db 358 N-----NDLTNATGNNISLLQVEGID 379
RESULT 15
US-09-206-942-63
; Sequence 63, Application US/09206942
; Patent No. 6432669
GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-ping
; APPLICANT: Kiehl, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; FILE REFERENCE: 1038-861 MIS:ib
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-63

Query Match 6.8%; Score 117; DB 4; Length 1188;
Best Local Similarity 22.3%; Pred. No. 0.14;
Matches 87; Conservative 49; Mismatches 143; Indels 112; Gaps 20;
QY 3 AADAPAGTDTIINQNTNVTG--IDSGTIVYVPHQAGYVKLVNYSVNSAVKGDFTFKITV 60
Db 37 SADTPKRNNTKTTLTNSTLEKILARGSFVNITANNEIRVNSDINI-----GGNSHLTL 90
QY 61 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYI 120
Db 91 WSKNKN-----SGVLINGNITSTAN--GNLTIYSSGWVDIHKNTLES-- 132
QY 121 DPENV-KKTGNVTLATGIGST-TANKTVLDVYKGYKFNLSIKCT-----IDQIDKT 171
Db 133 GLRLNTTREGDVAPEKGNLITGOGTITAGNNKGFRENVSLNGTGLFLNLSRPOKN 192
QY 172 N-----NTYRQTIYVNPSSD-NVIAPVLACN-----LKPNTDSNALIDQON 211
Db 193 NSLVTNYEGLNLSGVSNIEMIPNATSNWYSRYKGRTYWNITHLNASDSNF-----N 247
QY 212 TSIIKVKVDNADLSSEYFVNPENEDVTN-----SVNITFPNP-----NOYK----- 254
Db 248 LTIDSSAEDGSAPLLSSYTLNGISFTTDTTFNVNNAKVNFIKAPICTINOYNLNLVAL 307
QY 255 -----VEFTPDQITTPYVVVNGHIDPNSKGLALR-----STLYGYNS 295
Db 308 FNGNLSVSGGNNVTRLNASSNQOTPGVIINSKHLNA-SKGS-SLRETTGCTKVGFLI 365
QY 296 NIIWRMSWDNEAFNNGSGS-----GDGID 321
Db 366 N-----NDLTNATGNNISLLQVEGID 387

RESULT 16

US-09-002-285-96
; Sequence 96, Application US/09002285
; Patent No. 6369213
GENERAL INFORMATION:
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Wicker, Carol
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Walz, Michelle
; APPLICANT: Stockhoff, Brian
; APPLICANT: Muller-Cohn, Judy
; TITLE OF INVENTION: Toxins Active Against Pests
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/002,285
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/886,615
; FILING DATE: 1-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/674,002
; FILING DATE: 1-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-701C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 789 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-002-285-96

Query Match 6.7%; Score 116; DB 4; Length 789;
Best Local Similarity 22.0%; Pred. No. 0.091;
Matches 82; Conservative 57; Mismatches 132; Indels 102; Gaps 19;
QY 12 DITNQLTNTVTDIGSGTIVYVPHQAGYVKLVNYSVNSAVKGDFTFKITVPKELNLNGVTS 71
Db 148 EISOKLDIINVNLINSLTLEITPAYQRIKY-----VNEKFEELTFATETSSKVKKG--S 201
QY 72 TAKVPPIMAGDQVLANGVIDSDGNVIYFTDYVNTKDDV-----KATLTMPAYTDP 122
Db 202 PADILDELTELTELAKSVTKND---VDGFEFYLNTFHDVMVGNLFGRSALKRTASELITK 258
QY 123 ENVKKTGN-----VPLATIGS-----TTANKTV---LVDY-----EKYKGFY 157
Db 259 ENVTSGSEGVNPNFLIVLTAQAKAFLTUTTCRKLGLADIDVTSIMNEHLNKEKEEP 318
QY 158 NLSKTGTTDQDKTNTNTYRQTIYVNPSSDGNVIAPVLGTGNLKPNTDSNALI--DQONTSIK 215
Db 319 RVNILLPTL-----SNTFSNPNYAKVKGSDKAMIV-EAKP---GHALIGFELNSDSIT 368
QY 216 VYKVDNAADLSSEYFVN-----PENFEDVTNSVNTFPNPQY---KV 255
Db 369 VLKV-YEAKLKQYQVDRKDSLSEVIYDMDKLLCPDQSEQIYYTNNIVF--PNEYVITKI 425


```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 790 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 36a
US-09-073-898-4

Query Match
Best Local Similarity 6.6%; Score 115; DB 4; Length 790;
Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

QY 12 DITNLTNTVTVGIDSGTTPVPHQAGYVKLVNGFSVPNSAVKGDFTKITVPKELNGLVTS 71
DB 148 EISDKLDIINVNLINSLTTEITPAYQRIKY-----VNEKFELTFATETSSKVKKG--S 201
QY 72 TAKVPPIMAGDQVLANGVIDSGNVIYTFDYVNTKDDV-----KATLTWPAYIDP 122
DB 202 PANILDELTTELAKSVTKND---VDGFEFYLNTFHDVMVGNLFGRSALKTASELITK 258
QY 123 ENVKKTGN-----VTLATGIGS-----TTANKTV---LVDY-----EKYKGFY 157
DB 259 ENVKTSGSEGVNINFLIVLTAQAKAFLLTTCRKLGLADIDYVSTMNEHLNKEKEEF 318
QY 158 NLSIKGTIDQDKTNTYRTIYVNPNGDNVIAPVLTGNLKPNTDSNALI--DOQNTSIK 215
DB 319 RVNLTPTL-----SNTFSNPYAKVKGSDAKMIV-EAKP---GHALIGFEISNDSIT 368
QY 216 VYKVDNAADLSESYFN-----PENFEDVTNSVNIITFPNPQY---KV 255
DB 369 VLKV-YEAKLKQNYQVQDKDSLSEVIYGMKLLCPDQSEQIYYTNNIVF--PNEYVITKI 425
QY 256 EF---NTPDDQITTPYIVVYNGHIDPNKSGDLALRSTLYGYSNIWRSMSWNEVAFN 311
DB 426 DFTKKMKTLYEVTANFYDSSTGEIDLNKK-----KVESSEAEYK 465
QY 312 NGSGGDCIDKPV 324
DB 466 TLSANDDGVYMP 478

RESULT 19
US-08-471-033-29
; Sequence 29, Application US/08471033
; Patent No. 5770696
GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471.033
```

```
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: P-40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 790 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-471-033-29

Query Match
Best Local Similarity 6.6%; Score 114; DB 1; Length 789;
Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

QY 12 DITNLTNTVTVGIDSGTTPVPHQAGYVKLVNGFSVPNSAVKGDFTKITVPKELNGLVTS 71
DB 148 EISDKLDIINVNLINSLTTEITPAYQRIKY-----VNEKFELTFATETSSKVKKG--S 201
QY 72 TAKVPPIMAGDQVLANGVIDSGNVIYTFDYVNTKDDV-----KATLTWPAYIDP 122
DB 202 PADILDELTELAKSVTKND---VDGFEFYLNTFHDVMVGNLFGRSALKTASELITK 258
QY 123 ENVKKTGN-----VTLATGIGS-----TTANKTV---LVDY-----EKYKGFY 157
DB 259 ENVKTSGSEGVNINFLIVLTAQAKAFLLTTCRKLGLADIDYVSTMNEHLNKEKEEF 318
QY 158 NLSIKGTIDQDKTNTYRTIYVNPNGDNVIAPVLTGNLKPNTDSNALI--DOQNTSIK 215
DB 319 RVNLTPTL-----SNTFSNPYAKVKGSDAKMIV-EAKP---GHALIGFEISNDSIT 368
QY 216 VYKVDNAADLSESYFN-----PENFEDVTNSVNIITFPNPQY---KV 255
DB 369 VLKV-YEAKLKQNYQVQDKDSLSEVIYGMKLLCPDQSEQIYYTNNIVF--PNEYVITKI 425
QY 256 EF---NTPDDQITTPYIVVYNGHIDPNKSGDLALRSTLYGYSNIWRSMSWNEVAFN 311
DB 426 DFTKKMKTLYEVTANFYDSSTGEIDLNKK-----KVESSEAEYR 465

RESULT 20
US-08-471-044-29
; Sequence 29, Application US/08471044
; Patent No. 5840868
GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
```


TITLE OF INVENTION: No. 5840868e1 Pesticidal Proteins and Strains

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471,044

FILING DATE: 06-JUN-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/463,483

FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/314,594

FILING DATE: 09-SEP-1994

APPLICATION DATA:

APPLICATION NUMBER: US 08/218,018

FILING DATE: 23-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/037,057

FILING DATE: 25-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Pace, Gary M.

REGISTRATION NUMBER: 40,403

REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLV3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8582

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 789 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-471-044-29

Query Match 6.6%; Score 114; DB 2; Length 789;
Best Local Similarity 21.7%; Pred. No. 0.14;
Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

12 DITNLTNTVVGIDSGTTVYPHQAGYVKLVNFGVSPNSAVKGDTEKITVPKELNLNGVTS 71

148 EISDKLDIINVNLINSLTLEITPAYQRIKY---VNEKFEELTATETSSKVKKDG--S 201

72 TAKVPPIMAGDOVLANGVIDSGNVIYFTDVTNKKDDV-----KATLTMPAYIDP 122

202 PADILDELTELELAKSVTKND---VDGFEFLNTFHDVVMGNLFGRSALKTASELITK 258

123 ENVKKTGN-----VTLATIGS-----TTANKTV---LVDY-----EKYGRFY 157

259 ENVKTSGEVGNVNFVLVTLAQAFLLTTCRKLGLADIDYTSIMNEHLNKEKEEF 318

158 NLSIKGTIDQIDKNTYRQTIYVNPNSGDNVPIAPVLTKNPKNTDSNALI---DOQNTSIT 215

319 RVNILPTL-----SNTSPNPNYAKVGSDEDAKIV-EAKP---GHALIGEISNDSIT 368

216 VYKVDNAADLSSEYFVN-----PENFEDVNSVNIITFPNPQY---KV 255

369 VLKV-YEARLKQYQVQDKSLSEVIYGDMDKLLCPDQSEQIYVYTNVIF--PNEYVITKI 425

256 EF----NTPDDOITPIYVNVNGHIDPNKSGDLALRSTLYGNSNIWRWSWDNEVAFN 311

436 DFTKKMLRVEVTANFYDSSSTGEIDLNNK-----KVESSEAEYR 465

QY 312 NCGSGDGDIDKPV 324

Db 466 TLSANDDGVYMWPL 478

RESULT 21

US-08-463-483A-29

: Sequence 29, Application US/08463483A

: Patent No. 5849870

: GENERAL INFORMATION:

: APPLICANT: Warren, Gregory W

: APPLICANT: Koziel, Michael G

: APPLICANT: Mullins, Martha A

: APPLICANT: Nye, Gordon J

: APPLICANT: Carr, Brian

: APPLICANT: Desai, Ralini M

: APPLICANT: Kostichka, N. Kristy

: APPLICANT: Duck, Nicholas B

: APPLICANT: Estruch, Juan J

: TITLE OF INVENTION: No. 5849870e1 Pesticidal Proteins and Strains

: NUMBER OF SEQUENCES: 50

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: CIBA-GEIGY Corporation

: STREET: 7 Skyline Drive

: CITY: Hawthorne

: STATE: NY

: COUNTRY: USA

: ZIP: 10532

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent In Release #1.0, Version #1.30B

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/463,483A

: FILING DATE:

: CLASSIFICATION: 530

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 08/314,594

: FILING DATE: 09-SEP-1994

: APPLICATION DATA:

: APPLICATION NUMBER: US 08/218,018

: FILING DATE: 23-MAR-1994

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 08/037,057

: FILING DATE: 25-MAR-1993

: ATTORNEY/AGENT INFORMATION:

: NAME: Spruill, W. Murray

: REGISTRATION NUMBER: 32,943

: REFERENCE/DOCKET NUMBER: CGC 1695/CIP3

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 919-541-8615

: TELEFAX: 919-541-8689

: INFORMATION FOR SEQ ID NO: 29:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 789 amino acids

: TYPE: amino acid

: TOPOLOGY: linear

: MOLECULE TYPE: protein

: US-08-463-483A-29

Query Match 6.6%; Score 114; DB 2; Length 789;

Best Local Similarity 21.7%; Pred. No. 0.14;

Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

QY 12 DITNLTNTVVGIDSGTTVYPHQAGYVKLVNFGVSPNSAVKGDTEKITVPKELNLNGVTS 71

Db 148 EISDKLDIINVNLINSLTLEITPAYQRIKY---VNEKFEELTATETSSKVKKDG--S 201

QY 72 TAKVPPIMAGDOVLANGVIDSGNVIYFTDVTNKKDDV-----KATLTMPAYIDP 122

Db 202 PADILDELTELELAKSVTKND---VDGFEFLNTFHDVVMGNLFGRSALKTASELITK 258

```

QY 123 ENVKKTGN-----VTLATGIGS-----TTANKTV---LVDT-----EKYKFEY 157
Db 259 ENKVTSGSEGVNVLIVLTAQAQAFLLTTTCRKLGLADIDYTSIMNEHLNKEEF 318
QY 158 NLSIKGTTDIDKNTNTYRQTIYVNPNGDNIAPVLTGNLKPNTDSNALI--DQONTSTK 215
Db 319 RVNLIPTL-----SNTESNPNYAKVKGSEDAKIV-EAKP---GHALIGFEISNDSIT 368
QY 216 VYKVDNAADLSSEYFVN-----PENFEDVNTSNVITFPNPNOY---KV 255
Db 369 VLKV-YEAKLKQNYQVQKDSLSEVIYGDMLKLLCPDQSEQIYTYTNNIVF--PNEYVITKI 425
QY 256 EF---NTPDDQITPYIVVNGHIDPNKSGDLALRSTLYGYNISIIWRSMWDNEVAFN 311
Db 426 DFTKKMKTLRYEVNTANFYDSSTGEIDLNNK-----KVESSEAEYR 465
QY 312 NCGSGGDGIDKPV 324
Db 466 TLSANDDGVYMP 478

```

US-08-471-046A-29

Sequence 29, Application US/08471046A

Patent No. 5866326

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W

APPLICANT: Koziel, Michael G

APPLICANT: Mullins, Martha A

APPLICANT: Nye, Gordon J

APPLICANT: Carr, Brian

APPLICANT: Desai, Nalini M

APPLICANT: Kostichka, N. Kristy

APPLICANT: Duck, Nicholas B

APPLICANT: Estruch, Juan J

TITLE OF INVENTION: Method For Isolating Vegetative Insecticidal

TITLE OF INVENTION: Protein Genes

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5866326artis Corporation

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471,046A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/463,483

FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/314,594

FILING DATE: 09-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/218,018

FILING DATE: 23-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/037,057

FILING DATE: 25-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV8 - SOLv4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

```

; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 789 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-471-046A-29

```

Query Match 6.6%; Score 114; DB 2; Length 789;

Best Local Similarity 21.7%; Pred. No. 0.14;

Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

QY 12 DITNLTNTVTVGIDSGTTVYVPHQAGYVVKLVNGFVSPNSAVKGDTEKITVPKELNNGVTS 71

Db 148 EISDKLDIINVNLINSLTTEITPAYQRIKY-----VNEKFEELTETATETSSKVKKDG--S 201

QY 72 TAKVPPIMAGDOVLANGVIDSDGNVYTYTTOYVNTKDDV-----KATLTMPAYIDP 122

Db 202 PADILDELTELTELAKSVTKND---VDGFEFVLTDFHDMVGVNNGNLFGRSALKTASELITK 258

QY 123 ENVKKTGN-----VTLATGIGS-----TTANKTV---LVDT-----EKYKFEY 157

Db 259 ENKVTSGSEGVNVLIVLTAQAQAFLLTTTCRKLGLADIDYTSIMNEHLNKEEF 318

QY 158 NLSIKGTTDIDKNTNTYRQTIYVNPNGDNIAPVLTGNLKPNTDSNALI--DQONTSTK 215

Db 319 RVNLIPTL-----SNTESNPNYAKVKGSEDAKIV-EAKP---GHALIGFEISNDSIT 368

QY 216 VYKVDNAADLSSEYFVN-----PENFEDVNTSNVITFPNPNOY---KV 255

Db 369 VLKV-YEAKLKQNYQVQKDSLSEVIYGDMLKLLCPDQSEQIYTYTNNIVF--PNEYVITKI 425

QY 256 EF---NTPDDQITPYIVVNGHIDPNKSGDLALRSTLYGYNISIIWRSMWDNEVAFN 311

Db 426 DFTKKMKTLRYEVNTANFYDSSTGEIDLNNK-----KVESSEAEYR 465

QY 312 NCGSGGDGIDKPV 324

Db 466 TLSANDDGVYMP 478

RESULT 23

US-08-470-566B-29

Sequence 29, Application US/08470566B

Patent No. 5872212

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W

APPLICANT: Koziel, Michael G

APPLICANT: Mullins, Martha A

APPLICANT: Nye, Gordon J

APPLICANT: Carr, Brian

APPLICANT: Desai, Nalini M

APPLICANT: Kostichka, N. Kristy

APPLICANT: Duck, Nicholas B

APPLICANT: Estruch, Juan J

TITLE OF INVENTION: No. 5872212el Pesticidal Proteins and Strains

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5872212artis Corporation

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,566B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 530

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/463,483
 ; FILING DATE: 05-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/314,594
 ; FILING DATE: 09-SEP-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/218,018
 ; FILING DATE: 23-MAR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/037,057
 ; FILING DATE: 25-MAR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meigs, J. Timothy
 ; REGISTRATION NUMBER: 38,241
 ; REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV4 - SOLV4
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 919-541-8587
 ; TELEFAX: 919-541-8689
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 789 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-470-566B-29

Query Match 6.6%; Score 114; DB 2; Length 789;
 Best Local Similarity 21.7%; Pred. No. 0.14;
 Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;
 QY 12 DITNLTNTVYGVDSGTTVYHQAGYVKNLYGFSVPNSAVKGDTEKITVPKELNNGVTS 71
 Db 148 EISDKLDIINVNLINSLTETTPAYQRIKY---VNEKFEELTATETSSKVKKDG--S 201
 QY 72 TAKVPPIMAGDOVLANGVIDSGNVYFTDYVNTKDDV-----KATLTMPAYIDP 122
 Db 202 PADILDELTELAKSVTKND---VDGFEFLNTFHDVWVGNLFGRSALKTASELITK 258
 QY 123 ENVKKTGN-----VTLATIGS-----TTANKTV---LVQY-----EYGRFY 157
 Db 259 ENVKTSGEVGNVYFLVLTALQAQAFLLTTCRKLGLADIDVTSIMNEHLNKEKEEF 318
 QY 158 NLSIKGTIDQDKTNTYRTQTIYVNPNGDNVAPVLTGNLKNPTDSNALI--DQONTSIK 215
 Db 319 RVNLPRTL-----SNTFSPNPYAKVKGSDAKMIV-EAKP---GHALIGFEISNDSIT 368
 QY 216 VYKVDNAADLSSEYFVN-----PENFEDVTNSVNIITFPNPQY---KV 255
 Db 369 VLKV-YEAKLKONYQVQDKSLSEVIYGMKLLCPDQSEQIYTTNNIVF--PNEYVITKI 425
 QY 256 EF---NTPDDQITTPYIVVWNGHIDPNKSGDLALRSTLYGYNLSIWRSMDSMDNEVAFN 311
 Db 426 DFTKMKKTLRYEVTANFVDSSTGEIDLKK-----KVESSEAEYR 465
 QY 312 NGSGSGDGIDKRPV 324
 Db 466 TLSANDGVMPL 478

RESULT 24
 US-08-838-219B-2
 ; Sequence 2, Application US/08038219B
 ; Patent No. 5877012
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Gregory W
 ; APPLICANT: Koziel, Michael G
 ; APPLICANT: Mullins, Martha A
 ; APPLICANT: Nye, Gordon J
 ; APPLICANT: Carr, Brian
 ; APPLICANT: Desai, Ralini M
 ; APPLICANT: Kostichka, N. Kristy
 ; APPLICANT: Duck, Nicholas B

; APPLICANT: Estruch, Juan J
 ; TITLE OF INVENTION: A No. 5877012el Class of Proteins for the
 ; CONTROL OF PLANT PESTS
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CIBA-GEIGY Corporation
 ; STREET: 7 Skyline Drive
 ; CITY: Hawthorne
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10532
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30B
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/838,219B
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION NUMBER: US 08/463,483
 ; FILING DATE: 06-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/314,594
 ; FILING DATE: 09-SEP-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/218,018
 ; FILING DATE: 23-MAR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/037,057
 ; FILING DATE: 25-MAR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pace, Gary M.
 ; REGISTRATION NUMBER: 40,403
 ; REFERENCE/DOCKET NUMBER: CGC 1925
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 919-541-8582
 ; TELEFAX: 919-541-8689
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 789 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-838-219B-2

Query Match 6.6%; Score 114; DB 2; Length 789;
 Best Local Similarity 21.7%; Pred. No. 0.14;
 Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;
 QY 12 DITNLTNTVYGVDSGTTVYHQAGYVKNLYGFSVPNSAVKGDTEKITVPKELNNGVTS 71
 Db 148 EISDKLDIINVNLINSLTETTPAYQRIKY---VNEKFEELTATETSSKVKKDG--S 201
 QY 72 TAKVPPIMAGDOVLANGVIDSGNVYFTDYVNTKDDV-----KATLTMPAYIDP 122
 Db 202 PADILDELTELAKSVTKND---VDGFEFLNTFHDVWVGNLFGRSALKTASELITK 258
 QY 123 ENVKKTGN-----VTLATIGS-----TTANKTV---LVQY-----EYGRFY 157
 Db 259 ENVKTSGEVGNVYFLVLTALQAQAFLLTTCRKLGLADIDVTSIMNEHLNKEKEEF 318
 QY 158 NLSIKGTIDQDKTNTYRTQTIYVNPNGDNVAPVLTGNLKNPTDSNALI--DQONTSIK 215
 Db 319 RVNLPRTL-----SNTFSPNPYAKVKGSDAKMIV-EAKP---GHALIGFEISNDSIT 368
 QY 216 VYKVDNAADLSSEYFVN-----PENFEDVTNSVNIITFPNPQY---KV 255
 Db 369 VLKV-YEAKLKONYQVQDKSLSEVIYGMKLLCPDQSEQIYTTNNIVF--PNEYVITKI 425
 QY 256 EF---NTPDDQITTPYIVVWNGHIDPNKSGDLALRSTLYGYNLSIWRSMDSMDNEVAFN 311

Db 426 DFTKKMKTLRYEVTANFYDSSTGEIDLNNK-----KVESSEAEYR 465
QY 312 NCGSGSGDGDIDKPV 324
Db 466 TLSANDDGVYMP 478

RESULT 25
US-08-469-334-29
; Sequence 29, Application US/08469334
; Patent No. 5990383
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5990383el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,334
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/463,483
; FILING DATE:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8615
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 789 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-334-29

Query Match 6.6%; Score 114; DB 2; Length 789;
Best Local Similarity 21.7%; Pred. NO. 0.14;
Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;
QY 12 DITNLTNTVYGVDSGTTVYHQAGYVKLVNFGFVPSNAVKGDTFKITVPKELNLNGVTS 71
Db 148 EISDKLDIINVLNLTITPAIORIKY-----VNEKFEELTPATETSSKVKKDG--S 201

QY 72 TAKVPPIMAGQVLANGVIDSGNVIYTFDYVNTKDDV-----KATLTMPAYIDP 122
Db 202 PADILDELTELTELAKSVTKND---VDGFEFYLNTFHDVWVGNNLFGRSALKTASELITK 258
QY 123 ENVKKTGN-----VTLATGIGS-----TTANKTV---LVDY-----EKYGRFY 157
Db 259 ENVKTSGSEVGNVYNFLIVLTALQAQAFLLTTCRKLGLGLADIDVTSIMNEHLNKEKEF 318
QY 158 NLSIKGTIDQDKTNTNTYRQTIYVNPNSGDNVIAPVLTGNLKPNTDSNALI--DQONTSTK 215
Db 319 RVNILLPTL-----SNTFSNPNYAKVKGSDDEDAKMIV-EAKP---GHALIGEIFSNDST 368
QY 216 VYKVDNAADLSSEYFVN-----PENFEDVTNSVNIPTFPNNOY---KV 255
Db 369 VLKV-YEAKLKNYQVDKSLSEVIYGDMDKLLCPDQSEQIYYTTNNIVF--PNEVITIKI 425
QY 256 EF-----NTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWRMSMDNEVAFN 311
Db 426 DFTKKMKTLRYEVTANFYDSSTGEIDLNNK-----KVESSEAEYR 465
QY 312 NCGSGSGDGDIDKPV 324
Db 466 TLSANDDGVYMP 478

Search completed: June 23, 2003, 14:04:51
Job time : 22.367 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:03:36 ; Search time 46.9504 Seconds

(without alignments)
762.858 Million cell updates/sec

Title: US-10-056-052a-4

Perfect score: 1732

Sequence: 1 MVAADAPAAAGTDTNQLTNV.....NGSGGIDKPVVPEQPD 331

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues

1 number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

- Published_Applications_AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pdb.pdb.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pdb.pdb.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_PUB.pdb.pdb.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pdb.pdb.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pdb.pdb.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb.pdb.*
 - 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pdb.pdb.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUB.pdb.pdb.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pdb.pdb.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb.pdb.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pdb.pdb.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pdb.pdb.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pdb.pdb.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pdb.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1732	100.0	331	9	US-10-056-052-4
2	1732	100.0	345	10	US-09-813-820-7
3	1727	99.7	520	9	US-10-056-052-2
4	1723	99.5	936	1	US-08-781-986A-5249
5	1518	87.6	1021	10	US-09-815-242-5471
6	1518	87.6	1021	10	US-09-815-242-12544
7	409	23.6	767	10	US-09-815-242-5899
8	409	23.6	767	10	US-09-815-242-13140
9	346	20.0	978	10	US-09-815-242-5456
10	346	20.0	1001	10	US-09-815-242-12686
11	339	19.6	1018	10	US-09-815-242-5797
12	339	19.6	1018	10	US-09-815-242-12838
13	331	19.1	1027	1	US-08-781-986A-5254
14	282.5	16.3	251	1	US-08-781-986A-5252
15	257.5	14.9	1349	10	US-09-815-242-5898
16	257.5	14.9	1349	10	US-09-815-242-13137
17	232.5	13.4	932	10	US-09-815-242-5578
18	232.5	13.4	932	10	US-09-815-242-12438
19	210.5	12.2	841	10	US-09-815-242-5779

20	210.5	12.2	841	10	US-09-815-242-12751	Sequence 12751, A
21	136	7.9	559	1	US-08-781-986A-5251	Sequence 5251, Ap
22	130	7.5	900	9	US-09-884-465A-333	Sequence 333, App
23	130	7.5	906	9	US-09-884-465A-369	Sequence 369, App
24	130	7.5	906	9	US-09-884-465A-371	Sequence 371, App
25	130	7.5	906	9	US-09-884-465A-373	Sequence 373, App
26	129.5	7.5	1391	9	US-10-080-505-11	Sequence 11, Appl
27	125.5	7.2	921	9	US-09-117-447-6	Sequence 6, Appl
28	124.5	7.2	569	9	US-09-884-465A-235	Sequence 235, App
29	124.5	7.2	569	9	US-09-884-465A-255	Sequence 255, App
30	124.5	7.2	569	9	US-09-884-465A-256	Sequence 256, App
31	124.5	7.2	569	9	US-09-884-465A-356	Sequence 356, App
32	124.5	7.2	621	9	US-09-884-465A-368	Sequence 368, App
33	124.5	7.2	627	9	US-09-884-465A-366	Sequence 366, App
34	124.5	7.2	627	9	US-09-884-465A-367	Sequence 367, App
35	124.5	7.2	633	9	US-09-884-465A-349	Sequence 349, App
36	124.5	7.2	633	9	US-09-884-465A-350	Sequence 350, App
37	124.5	7.2	633	9	US-09-884-465A-351	Sequence 351, App
38	124.5	7.2	633	9	US-09-884-465A-352	Sequence 352, App
39	124.5	7.2	633	9	US-09-884-465A-353	Sequence 353, App
40	124.5	7.2	633	9	US-09-884-465A-354	Sequence 354, App
41	124.5	7.2	633	9	US-09-884-465A-355	Sequence 355, App
42	124.5	7.2	633	9	US-09-884-465A-357	Sequence 357, App
43	124.5	7.2	633	9	US-09-884-465A-358	Sequence 358, App
44	124.5	7.2	633	9	US-09-884-465A-359	Sequence 359, App
45	124.5	7.2	633	9	US-09-884-465A-360	Sequence 360, App
46	124.5	7.2	633	9	US-09-884-465A-361	Sequence 361, App
47	124.5	7.2	633	9	US-09-884-465A-362	Sequence 362, App
48	124.5	7.2	633	9	US-09-884-465A-363	Sequence 363, App
49	124.5	7.2	633	9	US-09-884-465A-364	Sequence 364, App
50	124.5	7.2	633	9	US-09-884-465A-365	Sequence 365, App
51	124.5	7.2	840	9	US-09-884-465A-10	Sequence 10, Appl
52	124.5	7.2	888	9	US-09-884-465A-338	Sequence 338, App
53	124.5	7.2	889	9	US-09-884-465A-348	Sequence 348, App
54	124.5	7.2	894	9	US-09-884-465A-336	Sequence 336, App
55	124.5	7.2	894	9	US-09-884-465A-337	Sequence 337, App
56	124.5	7.2	894	9	US-09-884-465A-340	Sequence 340, App
57	124.5	7.2	895	9	US-09-884-465A-342	Sequence 342, App
58	124.5	7.2	895	9	US-09-884-465A-344	Sequence 344, App
59	124.5	7.2	895	9	US-09-884-465A-346	Sequence 346, App
60	124.5	7.2	895	9	US-09-884-465A-347	Sequence 347, App
61	124.5	7.2	900	9	US-09-884-465A-334	Sequence 334, App
62	124.5	7.2	900	9	US-09-884-465A-335	Sequence 335, App
63	124.5	7.2	900	9	US-09-884-465A-339	Sequence 339, App
64	124.5	7.2	900	9	US-09-884-465A-341	Sequence 341, App
65	124.5	7.2	901	9	US-09-884-465A-343	Sequence 343, App
66	124.5	7.2	901	9	US-09-884-465A-345	Sequence 345, App
67	124.5	7.2	906	9	US-09-884-465A-332	Sequence 332, App
68	124.5	7.2	906	9	US-09-884-465A-370	Sequence 370, App
69	124.5	7.2	906	9	US-09-884-465A-372	Sequence 372, App
70	124.5	7.2	913	9	US-09-884-465A-384	Sequence 384, App
71	124.5	7.2	999	9	US-09-884-465A-376	Sequence 376, App
72	124.5	7.2	999	9	US-09-884-465A-377	Sequence 377, App
73	124.5	7.2	1039	9	US-09-884-465A-6	Sequence 6, Appl
74	124.5	7.2	1126	9	US-09-884-465A-383	Sequence 383, App
75	124.5	7.2	1139	9	US-09-884-465A-380	Sequence 380, App
76	124.5	7.2	1152	9	US-09-884-465A-379	Sequence 379, App
77	124.5	7.2	1238	9	US-09-884-465A-381	Sequence 381, App
78	124.5	7.2	1365	9	US-09-884-465A-382	Sequence 382, App
79	124.5	7.2	1378	9	US-09-884-465A-378	Sequence 378, App
80	124.5	7.2	1391	9	US-10-080-505-15	Sequence 15, Appl
81	123.5	7.1	1411	9	US-10-080-505-17	Sequence 17, Appl
82	121.5	7.0	5795	10	US-09-815-242-12610	Sequence 12610, A
83	118	6.8	789	9	US-10-099-285-94	Sequence 94, Appl
84	116	6.7	789	9	US-10-099-285-96	Sequence 96, Appl
85	116	6.7	839	9	US-10-099-285-96	Sequence 96, Appl
86	115	6.6	790	10	US-09-850-351A-4	Sequence 4, Appl
87	114	6.6	789	9	US-10-099-285-78	Sequence 78, Appl
88	114	6.6	789	9	US-10-099-285-100	Sequence 100, Appl
89	114	6.6	789	10	US-09-850-351A-8	Sequence 8, Appl
90	113.5	6.6	1228	9	US-09-117-447-2	Sequence 2, Appl
91	113	6.5	392	9	US-10-040-949A-46	Sequence 46, Appl
92	113	6.5	540	1	US-08-781-986A-5225	Sequence 5225, Ap

93 113 6.5 540 1 US-08-781-986A-5242 Sequence 5242, Ap
94 113 6.5 789 9 US-10-099-285-80 Sequence 80, Appl
95 113 6.5 789 9 US-10-099-285-90 Sequence 90, Appl
96 113 6.5 789 10 US-09-850-351A-6 Sequence 6, Appl
97 113 6.5 1394 9 US-09-839-996-2 Sequence 2, Appl
98 113 6.5 1394 9 US-10-080-505-2 Sequence 2, Appl
99 113 6.5 2478 10 US-09-815-242-5816 Sequence 5816, Ap
100 113 6.5 2478 10 US-09-815-242-12967 Sequence 12967, A

ALIGNMENTS

RESULT 1
US-10-056-052-4
: Sequence 4, Application US/10056052
: Publication No. US20030099656A1
: GENERAL INFORMATION:
: APPLICANT: HUTCHINS, Jeff T
: APPLICANT: DOMANSKI, Paul
: APPLICANT: PATEL, Pratiksha
: APPLICANT: HALL, Andrea
: TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
: FILE REFERENCE: P070690S04/BAS
: CURRENT APPLICATION NUMBER: US/10/056,052
: CURRENT FILING DATE: 2002-04-19
: PRIOR APPLICATION NUMBER: 60/308,116
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/298,413
: PRIOR FILING DATE: 2001-06-18
: PRIOR APPLICATION NUMBER: 60/274,611
: PRIOR FILING DATE: 2001-03-12
: PRIOR APPLICATION NUMBER: 60/264,072
: PRIOR FILING DATE: 2001-01-26
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 331
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
US-10-056-052-4

Query Match 100.0%; Score 1732; DB 9; Length 331;
Best Local Similarity 100.0%; Pred. No. 7.2e-126;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MVAADAPAAAGTDTTNTLVGIDSGTTVPHQAGYVKLVNPGFVSPNSAVKGDFTKTV 60
1 MVAADAPAAAGTDTTNTLVGIDSGTTVPHQAGYVKLVNPGFVSPNSAVKGDFTKTV 60
61 PKELNLCVSTAKVPPIMAGDQVLAVGIDSGNVITFTDYVNTKDDVATLTMPAYI 120
61 PKELNLCVSTAKVPPIMAGDQVLAVGIDSGNVITFTDYVNTKDDVATLTMPAYI 120
121 DPENVKKTGNTLVATIGSTTANKTTLVDYKYGKFNLSIKGTIDQDKTNTNTYRTIY 180
121 DPENVKKTGNTLVATIGSTTANKTTLVDYKYGKFNLSIKGTIDQDKTNTNTYRTIY 180
181 VNPAGDNIAPVLGTLNPKNTDSNALIDQNTSTKVKYKVDNAALDSESYFVNPENFEDVT 240
181 VNPAGDNIAPVLGTLNPKNTDSNALIDQNTSTKVKYKVDNAALDSESYFVNPENFEDVT 240
241 NSVNTTFPNQYKVEFNTPDQITTPYIVVNVGHIDPNSKGLALRSTLYGNSNIWR 300
241 NSVNTTFPNQYKVEFNTPDQITTPYIVVNVGHIDPNSKGLALRSTLYGNSNIWR 300
301 SMSMDNEVAFNNGSGDGDIDKPVVPEQPD 331
301 SMSMDNEVAFNNGSGDGDIDKPVVPEQPD 331

RESULT 2

US-09-813-820-7
: Sequence 7, Application US/09813820
: Patent No. US20020102262A1
: GENERAL INFORMATION:
: APPLICANT: Hook, Magnus
: Patti, Joseph M.
: House-Pompeo, Karen
: Sthanam, Narayana
: Symersky, Jindrich
: TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
: AND METHODS OF USE
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: Texas
: COUNTRY: U.S.
: ZIP: 77210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/813,820
: FILING DATE: 22-Mar-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/856,253
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Kitchell, Barbara S.
: REGISTRATION NUMBER: 33,928
: REFERENCE/DOCKET NUMBER: TAMK:193
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (512) 474-7577
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 345 amino acids
: TYPE: amino acid
: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-813-820-7

Query Match 100.0%; Score 1732; DB 10; Length 345;
Best Local Similarity 100.0%; Pred. No. 7.6e-126;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVAADAPAAAGTDTTNTLVGIDSGTTVPHQAGYVKLVNPGFVSPNSAVKGDFTKTV 60
DB 13 MVAADAPAAAGTDTTNTLVGIDSGTTVPHQAGYVKLVNPGFVSPNSAVKGDFTKTV 72
QY 61 PKELNLCVSTAKVPPIMAGDQVLAVGIDSGNVITFTDYVNTKDDVATLTMPAYI 120
DB 73 PKELNLCVSTAKVPPIMAGDQVLAVGIDSGNVITFTDYVNTKDDVATLTMPAYI 132
QY 121 DPENVKKTGNTLVATIGSTTANKTTLVDYKYGKFNLSIKGTIDQDKTNTNTYRTIY 180
DB 133 DPENVKKTGNTLVATIGSTTANKTTLVDYKYGKFNLSIKGTIDQDKTNTNTYRTIY 192
QY 181 VNPAGDNIAPVLGTLNPKNTDSNALIDQNTSTKVKYKVDNAALDSESYFVNPENFEDVT 240
DB 193 VNPAGDNIAPVLGTLNPKNTDSNALIDQNTSTKVKYKVDNAALDSESYFVNPENFEDVT 252
QY 241 NSVNTTFPNQYKVEFNTPDQITTPYIVVNVGHIDPNSKGLALRSTLYGNSNIWR 300
DB 253 NSVNTTFPNQYKVEFNTPDQITTPYIVVNVGHIDPNSKGLALRSTLYGNSNIWR 312
QY 301 SMSMDNEVAFNNGSGDGDIDKPVVPEQPD 331
|||||

Db 313 SMSWDNEAFNNGSGGDIKPKVPVPEQDE 343

RESULT 3

US-10-056-052-2
Sequence 2, Application US/10056052
Publication NO. US20030099656A1
GENERAL INFORMATION:
APPLICANT: HUTCHINS, Joseph M
APPLICANT: PATTI, Jeff T
APPLICANT: DOMANSKI, Paul
APPLICANT: PATEL, Pratiksha
APPLICANT: HALL, Andrea
TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
FILE REFERENCE: P07069US04/BAS
CURRENT APPLICATION NUMBER: US/10/056,052
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/308,116
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/298,413
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/274,611
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/264,072
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 520
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-056-052-2

Query Match 99.7%; Score 1727; DB 9; Length 520;
Best Local Similarity 100.0%; Pred. No. 3.2e-125; Mismatches 0; Indels 0; Gaps 0;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 VAADAPAGTDTITNLTNTVTVGIDSGTTVPYHQAGYVKLVNLYGFSVPNSAVKGDTEKITVP 61
Db 182 VAADAPAGTDTITNLTNTVTVGIDSGTTVPYHQAGYVKLVNLYGFSVPNSAVKGDTEKITVP 241
QY 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDVTNKKDDVKATLTMPAYID 121
Db 242 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDVTNKKDDVKATLTMPAYID 301
QY 122 PENVKKTGNVTATGIGSTTANKTVLDYKGYKFNLSIKGTIDQIDKTNNTYRQTIYV 181
Db 302 PENVKKTGNVTATGIGSTTANKTVLDYKGYKFNLSIKGTIDQIDKTNNTYRQTIYV 361
QY 182 NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSESYFVNPENFEDVTN 241
Db 362 NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSESYFVNPENFEDVTN 421
QY 242 SVNITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGNSNIWRS 301
Db 422 SVNITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGNSNIWRS 481
QY 302 MSWDNEAFNNGSGGDIKPKVPVPEQDE 331
Db 482 MSWDNEAFNNGSGGDIKPKVPVPEQDE 511

RESULT 4

US-08-781-986A-5249
Sequence 3249, Application US/08781986A
Publication NO. US20030054436A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue

CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5249

Query Match 99.5%; Score 1723; DB 1; Length 936;
Best Local Similarity 99.7%; Pred. No. 1.4e-124; Mismatches 1; Indels 0; Gaps 0;
Matches 329; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 VAADAPAGTDTITNLTNTVTVGIDSGTTVPYHQAGYVKLVNLYGFSVPNSAVKGDTEKITVP 61
Db 230 VAADAPAGTDTITNLTNTVTVGIDSGTTVPYHQAGYVKLVNLYGFSVPNSAVKGDTEKITVP 289
QY 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDVTNKKDDVKATLTMPAYID 121
Db 290 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDVTNKKDDVKATLTMPAYID 349
QY 122 PENVKKTGNVTATGIGSTTANKTVLDYKGYKFNLSIKGTIDQIDKTNNTYRQTIYV 181
Db 350 PENVKKTGNVTATGIGSTTANKTVLDYKGYKFNLSIKGTIDQIDKTNNTYRQTIYV 409
QY 182 NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSESYFVNPENFEDVTN 241
Db 410 NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSESYFVNPENFEDVTN 469
QY 242 SVNITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGNSNIWRS 301
Db 470 SVNITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGNSNIWRS 529
QY 302 MSWDNEAFNNGSGGDIKPKVPVPEQDE 331
Db 530 MSWDNEAFNNGSGGDIKPKVPVPEQDE 559

RESULT 5

US-09-815-242-5471
Sequence 5471, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5471
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5471

Query Match 87.6%; Score 1518; DB 10; Length 1021;
Best Local Similarity 87.3%; Pred. No. 1e-108;
Matches 288; Conservative 17; Mismatches 25; Indels 0; Gaps 0;
QY 2 VAADAPAGTDTITNLTNTVTVGIDSGTTVYPHQAGYVKLVNKGFSVPNSAVKGDTEKITVP 61
DB 221 VAADAPAGTDTITNLTNTVTVGIDSGTTVYPHQAGYVKLVNKGFSVPNSAVKGDTEKITVP 280
QY 62 KELNLNGVTSTAKVPPIMVGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 121
DB 281 KELNLNGVTSTAKVPPIMVGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 340
QY 122 PENVKKTGNVTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKNTNTYRQTIVV 181
DB 341 PENVKKTGNVTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKNTNTYRQTIVV 400
QY 182 NPSGDNVLPALVTGNLKPNTSDNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 241
DB 401 NPSGDNVLPALVTGNLKPNTSDNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 460
QY 242 SVNITFPNPNQYKVEFNTPDQDITTPYIVVNGHIDPNSKGDALRLSTLYGYNSTIWR 301
DB 461 QVRISFPNPNQYKVEFNTPDQDITTPYIVVNGHIDPNSKGDALRLSTLYGYNSTIWR 520
QY 302 MSWDNEVAFNNGSGSDGIDKPVVPEQDPE 331
DB 521 MSWDNEVAFNNGSGSDGIDKPVVPEQDPE 550

RESULT 6
US-09-815-242-12544
; Sequence 12544, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12544
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12544

Query Match 87.6%; Score 1518; DB 10; Length 1021;
Best Local Similarity 87.3%; Pred. No. 1e-108;
Matches 288; Conservative 17; Mismatches 25; Indels 0; Gaps 0;
QY 2 VAADAPAGTDTITNLTNTVTVGIDSGTTVYPHQAGYVKLVNKGFSVPNSAVKGDTEKITVP 61
DB 221 VAADAPAGTDTITNLTNTVTVGIDSGTTVYPHQAGYVKLVNKGFSVPNSAVKGDTEKITVP 280
QY 62 KELNLNGVTSTAKVPPIMVGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 121
DB 281 KELNLNGVTSTAKVPPIMVGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 340
QY 122 PENVKKTGNVTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKNTNTYRQTIVV 181
DB 341 PENVKKTGNVTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKNTNTYRQTIVV 400
QY 182 NPSGDNVLPALVTGNLKPNTSDNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 241
DB 401 NPSGDNVLPALVTGNLKPNTSDNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 460
QY 242 SVNITFPNPNQYKVEFNTPDQDITTPYIVVNGHIDPNSKGDALRLSTLYGYNSTIWR 301
DB 461 QVRISFPNPNQYKVEFNTPDQDITTPYIVVNGHIDPNSKGDALRLSTLYGYNSTIWR 520
QY 302 MSWDNEVAFNNGSGSDGIDKPVVPEQDPE 331
DB 521 MSWDNEVAFNNGSGSDGIDKPVVPEQDPE 550

RESULT 7
US-09-815-242-5899
; Sequence 5899, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21


```
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5899
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5899

Query Match      23.6%  Score 409;  DB 10;  Length 767;
Best Local Similarity 28.9%;  Pred. No. 2.1e-23;
Matches 101;  Conservative 72;  Mismatches 131;  Indels 46;  Gaps 11;

Qy  2 VAADAPAAAGTDTITNQLTNVTGIDSGT-----TVYPHOAGYVKLVNFGFSPNSAV 51
Db  155 VVEETKATGTDVTNK-----VEVEEGSEIVGHKQDTNVVNPNAERVTLYKKWFGEGIK 209

Qy  52 KGDTFKITVPKELNLNGVTSTAKYPPIMAGD-QVLANGVIDSDGNVIYTFDYVNTKDDV 110
Db  210 AGDYFDFTLSDNVETHGISTLRKVPKISTDQGYMATGEIIGERKRVYTFKYYVQEKD 269

Qy  111 KATLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVVDY-----EKYGRFYNLSIKGTI 165
Db  270 TAEISLNLFDIPTVTQKGNQNVKLGETTIVSKIFNIQYLGVRDNG-----VTANGRI 325

Qy  166 DQIDKTNNTYQTITVYVNPNGDNVIAPIVLTGNL-----KPTDSNALIDQNTSIKVKYVDN 221
Db  326 DTLNKGDKGSHFAYMKPNQSLSSVTVTGQVTKGNKPGVN-----NPTVVKYKHIG 377

Qy  222 AADLSESYFVNPN-----FEDVTSNVIITFPNPNQYKVEFNTPDDQITPTIYVVVNGHIDP 278
Db  378 SDDLAEVYAKLDDVSKFEDVTDNMSLDFDTNGGYSFLNFNLDQ--SKNYVIKYEYGYDS 435

Qy  279 NSKGDALRLSTLYGYNISNIWRSMWDNEVAFNNGSGSGDGIDK---PVV 325
Db  436 NA-SNLEFQTHLFGYNYNYTSLNLTWKNGVAFYSNNAAGDGKDKLKEPII 484

RESULT 8
US-09-815-242-13140
; Sequence 13140, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5899
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5899
```

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13140
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13140

Query Match      23.6%  Score 409;  DB 10;  Length 767;
Best Local Similarity 28.9%;  Pred. No. 2.1e-23;
Matches 101;  Conservative 72;  Mismatches 131;  Indels 46;  Gaps 11;

Qy  2 VAADAPAAAGTDTITNQLTNVTGIDSGT-----TVYPHOAGYVKLVNFGFSPNSAV 51
Db  155 VVEETKATGTDVTNK-----VEVEEGSEIVGHKQDTNVVNPNAERVTLYKKWFGEGIK 209

Qy  52 KGDTFKITVPKELNLNGVTSTAKYPPIMAGD-QVLANGVIDSDGNVIYTFDYVNTKDDV 110
Db  210 AGDYFDFTLSDNVETHGISTLRKVPKISTDQGYMATGEIIGERKRVYTFKYYVQEKD 269

Qy  111 KATLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVVDY-----EKYGRFYNLSIKGTI 165
Db  270 TAEISLNLFDIPTVTQKGNQNVKLGETTIVSKIFNIQYLGVRDNG-----VTANGRI 325

Qy  166 DQIDKTNNTYQTITVYVNPNGDNVIAPIVLTGNL-----KPTDSNALIDQNTSIKVKYVDN 221
Db  326 DTLNKGDKGSHFAYMKPNQSLSSVTVTGQVTKGNKPGVN-----NPTVVKYKHIG 377

Qy  222 AADLSESYFVNPN-----FEDVTSNVIITFPNPNQYKVEFNTPDDQITPTIYVVVNGHIDP 278
Db  378 SDDLAEVYAKLDDVSKFEDVTDNMSLDFDTNGGYSFLNFNLDQ--SKNYVIKYEYGYDS 435

Qy  279 NSKGDALRLSTLYGYNISNIWRSMWDNEVAFNNGSGSGDGIDK---PVV 325
Db  436 NA-SNLEFQTHLFGYNYNYTSLNLTWKNGVAFYSNNAAGDGKDKLKEPII 484

RESULT 9
US-09-815-242-5456
; Sequence 5456, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
```

Query Match 20.0%; Score 346; DB 10; Length 1001;
Best Local Similarity 27.2%; Pred. No. 2.1e-18;
Matches 91. Conservative 69. Mismatches 148. Indels 26

Query Match	20.0%	Score	346;	DB	10;	Length	1001;
Best Local Similarity	27.2%	Pred. No.	2.le-18;				
Matches	91;	Conservative	69;	Mismatches	148;	Indels	26;
				Gaps	10;		
Qy	3	AADAPA-AGTDITNLTNTVTVGIDS--GTTVYPHQAGVVKVKNLYGFSVPNSAVKGDTRFIT	59				
Db	186	ASDAKVETGDTVTSKVTVESGSIIEAPQGNKVPEHAGORVWLJKLKFEGUHKHGDYFDT	245				
Qy	60	VPEKLNELNGVTSYAKVPPIMAGDVLANGVIDSDGNVYTFDVTXNTKDDYKATLTMPAY	119				
Db	246	LSNNVNTYGVSTARKVPEIKNGSVVMATGQLLNGKIRYTFDVIDYKVNVTADLEINLF	305				
Qy	120	IDPENVKKTGNVTLATGIGSTTANKTVLVDYKEYGKFNLSIKGTIDOIDKTNNTYRQTI	179				
Db	306	IDPKTVQSGQQTITSTLNDKETNTLPYIEYNGPVSNSYANVNGSIETFDKGNRRTHVA	365				
Qy	180	YVNP-SCDNVIAPVLTGNTLNPNTSDNALIDQONTISIKVYKV-DNAADLSSEYFVNPN--	235				
Db	366	YIKPQNGHKSDSVISITGLTQGSKA----DGKAPTVMKYVEVLKDNAPLQPSVYANVSDSS	421				
Qy	236	-FEDVTNSV--NTTFPMPNOVKVFENPTDQITTPYIVVNGHIDPNKSGDLALRSTLYG	292				
Db	422	MFDVTEEMKDKLVENNGNYKKLDI----EKLEKSYVIHYDGEYLSGS-DQVNFRTIMFG	476				
Qy	293	YNS-----NIIWRMSWDNEVAFNNGSGSDG	319				
Db	477	YPEOQKYYVTHLYGKYLTDPDGLVLYSNKAKGDG	510				

```

RESULT 11
US-09-815-242-5797
; Sequence 5797, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Hazelbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/131,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5797
; LENGTH: 1018

```


Db 345 DLSRN---MTVYNQPKKTYTKETFTVNLTKYKFNPAK-----NFKIYEVTNQNF 393
Qy 226 SESYFVNPFEDVTSVNTTFPNQ-YKVEFTPDQITPIVVVNGHIDPNSKGD 284
Db 394 VDSFTPTDTSKLTVDTKFKITYSDNKTATVDLLNGSSDKQYIIQQVAYPDNSSTDNG 453
Qy 285 ALRSTLYGYNISWISWSDNEVAFNNGSGDGD 322
Db 454 KIDYLTETQNG-----KSSWSNSYSNVNGSGSTANGDOK 486

RESULT 18

US-09-815-242-12438

; Sequence 12438, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12438

; LENGTH: 932

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

; US-815-242-12438

; Query Match 13.48; Score 232.5; DB 10; Length 932;

; Best Local Similarity 24.98; Pred. No. 1.1e-09;

; Matches 84; Conservative 55; Mismatches 158; Indels 41; Gaps 12;

Qy 6 APAAGTDTITNQL--TNVTGIDSG-----TTVPHQAGYKVLNKGFSVPNSAVKGD 55

Db 169 APOGTVNVDKVFHTNDIAIDKGVNKTGTNTEFWATSSDVLKLNKANYTIDSVKREGDT 228

Qy 56 FKITVPKELNLTSTAKVPPIM-AGDQVLANGVIDSD-GNVIYTFDVTNKKDDVKAT 113

Db 229 FTKYGYQYFRPGSVRLPSQTNLYNAOQNIAGIYDSTSTTTTFTNVDQYTNISGS 288

Qy 114 LTMPAYIDPEN--VKKTG---NVTLATIGSTTANKTVLDYKYGKFNLSIKGTIDQI 168

Db 289 FEQVAFKRNATTDKTAYPEVTL-----GNDYKSNVIYDGNKQKQQLISSTNYNNE 344

Qy 169 DKTNTYRQIYVNPNSGDNVIAPIVLTGNL---KPNTDSNALIDQNTSIKVKYVDNAADL 225

Db 345 DLSRN---MTVYNQPKKTYTKETFTVNLTKYKFNPAK-----NFKIYEVTNQNF 393

Qy 226 SESYFVNPFEDVTSVNTTFPNQ-YKVEFTPDQITPIVVVNGHIDPNSKGD 284

Db 394 VDSFTPTDTSKLTVDTKFKITYSDNKTATVDLLNGSSDKQYIIQQVAYPDNSSTDNG 453
Qy 285 ALRSTLYGYNISWISWSDNEVAFNNGSGDGD 322
Db 454 KIDYLTETQNG-----KSSWSNSYSNVNGSGSTANGDOK 486

RESULT 19

US-09-815-242-5779

; Sequence 5779, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5779

; LENGTH: 841

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

; US-09-815-242-5779

; Query Match 12.28; Score 210.5; DB 10; Length 841;

; Best Local Similarity 23.18; Pred. No. 4.7e-08;

; Matches 83; Conservative 57; Mismatches 136; Indels 83; Gaps 16;

Qy 6 APAAGTDTITNQL--TNVTGIDSG-----TTVPHQAGYKVLNKGFSVPNSAVKGD 55

Db 178 APOGTVNVDKVFHTNDIAIDKGVNKTGTNTEFWATSSDVLKLNKANYTIDSVKREGDT 237

Qy 56 FKITVPKELNLTSTAKVPPIM-AGDQVLANGVIDSDGN-VIYTFDVTNKKDDVKAT 113

Db 238 FTKYGYQYFRPGSVRLPSQTNLYNAOQNIAGIYDSTTTTFTNVDQYTNVRS 297

Qy 114 LTMPAYIDPEN--VKKTG---NVTLATIGSTTANKTVLDYKYGKFNLSIKGTIDQI 168

Db 298 FEQVAFKRNATTDKTAYPEVTL-----GNDYSEIIVDYG-----NKKAOPLISST 347

Qy 169 DKTNN---TVRQIYVNPNSGDNVIAPIVLTGNL-----KPNTDSNALIDQNTSIKVKVD 220

Db 348 NYINNEDELNSMTAYVNPQNPKYTKQTFVTLNLTGYKFNPA-----KNFKIYEVT 397

Qy 221 NAADLSSEYFVNPFEDVTSVNTTFPNP-----QYKV-EPNTPDDQ 263

Db 398 DQNOQFVDSFTPDTSKLDKVDQDFDVIYSDNKTATVDLMKGOTSSNKQYIIQQVAYPDNS 457

Qy 264 ITTPIYVVNGHIDPNSKGDALRSTLYGYNISWISWSDNEVAFNNGSGDGD 322

Db 458 STD-----NGKIDYTLDTD-----KTKY-----SWSNSYSNVNGSSSTANGDOK 495

RESULT 20

US-09-815-242-12751

Sequence 12751, Application US/09815242

Patent No. US20020061589A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes In

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12751

LENGTH: 841

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-09-815-242-12751

Query Match 12.2%; Score 210.5; DB 10; Length 841;

Best Local Similarity 23.1%; Pred. No. 4.7e-08;

Matches 83; Conservative 57; Mismatches 136; Indels 83; Gaps 16;

6 APAAGTDTITNOL--TNVTVGIDSG-----TTVYPHQAQYVVKLVNGFSPNSAVKGD 55

178 APOQGTNVNDKRVHFSNIDIAIDKGHVNOTTKTEFWATSDVLKLNKANTIDDSVKEGDT 237

56 FKITVPKELNLNGVTSTAKVPPIM-AGDQVLANGVIDSDGN-VIYTFDYVNTKDDVKAT 113

238 FTEFKYGYFRGSPVRLPSQTNLYNAQGNIAKGIYDSTNTTYTFTYVQDQTNVRGS 297

114 LTMPAYIDPEN--VKKTG---NVLATGIGSTANKTVLVYDYKYGKFNLSIKGTIDQI 168

298 FEQVAFARKNATTDKATYKMEVTL-----GNDTYSEEIIVDYG-----NKKAOPLISST 347

169 DKTNN---TYRQTIYVNSGDNVAPVLGNL-----KPNSTDNALIDQNTSIKVKVD 220

348 NYINNEDLSRNMATVNPQKNTYTKQTEVTNLTKYKFNPN-----KNFKIYEVT 397

221 NAAALSESFYFNPENFEDVNTSVNITFPNP-----QYKV-EFNTPDQ 263

398 DONQFVDSFTPTDSKLKDTQDFDIYNDNKNKTATVLMKQTSNKNQYIIQQVAYPNDS 457

264 ITTPVIVVNGHIDPNKGDALRSTLYGYSNIIWRSMNSMDNEVAFNNGSGSGDGIDK 322

458 STD-----NGKIDYTLDTD-----KTKY-----SWSNSYSNVNGSSSTANGDOK 495

RESULT 21

US-08-781-986A-5251

Sequence 5251, Application US/08781986A

Publication No. US20030054436A1

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781.986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 5251:

SEQUENCE CHARACTERISTICS:

LENGTH: 559 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-781-986A-5251

Query Match 7.9%; Score 136; DB 1; Length 559;

Best Local Similarity 31.2%; Pred. No. 0.015;

Matches 29; Conservative 23; Mismatches 35; Indels 6; Gaps 3;

QY 236 FEDVTSNVITFPNPQYKVFENTPDDQITTPYIVVNGHIDPNKGDALRSTLYGYSN 295

Db 14 FEDVTNNLSLDFDTNGGYSLNPNLDQ--SKNYIVKIEGYGYSNA-SNLEFQTHLFGYIN 70

QY 296 NIIWRSMNSMDNEVAFNNGSGSGDGIDK---PVV 325

Db 71 YYTSLNLTWKNGVAFYSNNAQDGCKDKLKEPII 103

RESULT 22

US-09-884-465A-333

Sequence 333, Application US/09884465A

Publication No. US20030077293A1

GENERAL INFORMATION:

APPLICANT: Shire Blochem, Inc.

APPLICANT: Brodeur, Bernard

APPLICANT: Hamel, Josee

APPLICANT: Martin, Denis

APPLICANT: Charland, Nathalie

APPLICANT: Ouellet, Catherine

TITLE OF INVENTION: Streptococcus Antigens

FILE REFERENCE: 055190-0044

CURRENT APPLICATION NUMBER: US/09/884,465A

CURRENT FILING DATE: 2001-06-20

PRIOR APPLICATION NUMBER: 60/212,683

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 384

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 333
; LENGTH: 900
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
US-09-884-465A-333

Query Match.              7.5%; Score 130; DB 9; Length 900;
Best Local Similarity 23.6%; Pred. No. 0.083;
Matches 95; Conservative 47; Mismatches 152; Indels 108; Gaps 23;

QY 1 MVAADAPAAAGTDTITNOL--TNVTGIDSGTIVYPHOAG--YVKLNYGFSVPNSA----- 50
Db 279 IASKDYPEVSYDGTFTVPTSLEYKASQTIFFPHAGDTYLRVNPQFAVPKGTDLVRVF 338
QY 51 -----VKGDTPKITVPKELNNGVTSTA--KVPPIMAGDQVLANGVIDSDGNV 96
Db 339 DEFHGNAYLENNYKVGEIKLPIPK-LN-QGTRTAGNKIPVTF-----MANAYLDNQSTY 391
QY 97 IYTF-----TD-----YVNTKDDVKATLTMPA-YIDPENVKKTGNVTLA 134
Db 392 IVEVPILKEKQNTDRKPSILPQFKRKAQENSKLDEKVEEPTSEKVEKEKLSGTSNST-- 449
QY 135 TGIGSTTANKTVLVD--YEKYGK-----YNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNV- 188
Db 450 ---SNSTLEEVTPDPVQEKVAKFAESYGMKLENVLFNMDGTIELYL-----PSGEVIK 500
QY 189 -----IAPVLTGNLKPNTD---SNALIDQNTSIKVIKYVDNAADLSESYFVNPNFE 237
Db 501 KKNMADFTGEAPQNGENKPSKNGKYSTGTVENQPTENK--PADSLPEAPNEKPVKPE--- 555
QY 238 DVNTSNVITFPNPQYKVEFNTPDQITPTPIYVYVNGHIDP-----NSKGDIALRST 289
Db 556 ---NSTDNGMLNP-----EGNVGSDPMLDPALEEAAP-VDPVQEKLEKFTASYGLGLDSV 606
QY 290 LYGNSNIIRWSMWDNEVAFNNGSGSGDGIDKPVVPEQDE 331
Db 607 IFNMDGTIELRLPS--GEVIKKNLS---DFAIGPQITTYDDE 643

RESULT 24
US-09-884-465A-371
; Sequence 371, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 371
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
US-09-884-465A-371

Query Match              7.5%; Score 130; DB 9; Length 906;
Best Local Similarity 23.6%; Pred. No. 0.083;
Matches 95; Conservative 47; Mismatches 152; Indels 108; Gaps 23;

QY 1 MVAADAPAAAGTDTITNOL--TNVTGIDSGTIVYPHOAG--YVKLNYGFSVPNSA----- 50
Db 279 IASKDYPEVSYDGTFTVPTSLEYKASQTIFFPHAGDTYLRVNPQFAVPKGTDLVRVF 338
QY 51 -----VKGDTPKITVPKELNNGVTSTA--KVPPIMAGDQVLANGVIDSDGNV 96
Db 339 DEFHGNAYLENNYKVGEIKLPIPK-LN-QGTRTAGNKIPVTF-----MANAYLDNQSTY 391
QY 97 IYTF-----TD-----YVNTKDDVKATLTMPA-YIDPENVKKTGNVTLA 134
Db 392 IVEVPILKEKQNTDRKPSILPQFKRKAQENSKLDEKVEEPTSEKVEKEKLSGTSNST-- 449
QY 135 TGIGSTTANKTVLVD--YEKYGK-----YNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNV- 188
Db 450 ---SNSTLEEVTPDPVQEKVAKFAESYGMKLENVLFNMDGTIELYL-----PSGEVIK 500
QY 189 -----IAPVLTGNLKPNTD---SNALIDQNTSIKVIKYVDNAADLSESYFVNPNFE 237
Db 501 KKNMADFTGEAPQNGENKPSKNGKYSTGTVENQPTENK--PADSLPEAPNEKPVKPE--- 555
QY 238 DVNTSNVITFPNPQYKVEFNTPDQITPTPIYVYVNGHIDP-----NSKGDIALRST 289
Db 556 ---NSTDNGMLNP-----EGNVGSDPMLDPALEEAAP-VDPVQEKLEKFTASYGLGLDSV 606
QY 290 LYGNSNIIRWSMWDNEVAFNNGSGSGDGIDKPVVPEQDE 331
Db 607 IFNMDGTIELRLPS--GEVIKKNLS---DFAIGPQITTYDDE 643

RESULT 23
US-09-884-465A-369
; Sequence 369, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 369
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
US-09-884-465A-369

Query Match              7.5%; Score 130; DB 9; Length 906;
Best Local Similarity 23.6%; Pred. No. 0.083;
Matches 95; Conservative 47; Mismatches 152; Indels 108; Gaps 23;
```


Db 450 ---SNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYL-----PSGEVIK 500
Qy 189 -----IAPVLTGNLKPNTD---SNALIDQONTSIKVKVDNAADLSSESYFVNPNFE 237
Db 501 KNAADFTGEAPOGNGENKPSKGVSTGTVENQPTENK---PADSLPEAPNEKPVKPE--- 555
Qy 238 DVTNSVNIPTPNPNQYKVFENTPDDQITTPYIVVNGHIDP-----NSKGDALALRST 289
Db 556 ---NSTDNGMLNP-----EGNVGSDPMLDPALEAPA-VDPVQEKLEKFTASYGLGLDSV 606
Qy 290 LYGNSNIWIRSWDNEVAFNNGSGSGDGDIDRPVPEQDPE 331
Db 607 IFNMDGTIELRLPS--GEVIKKNLS---DFIAGPQITYTDD 643

RESULT 25

US-09-884-465A-373
Sequence 373, Application US/09884465A
Publication No. US20030077293A1
GENERAL INFORMATION:
APPLICANT: Shire Biochem, Inc.
APPLICANT: Hamel, Josee
APPLICANT: Brodeur, Bernard
APPLICANT: Martin, Denis
APPLICANT: Charland, Nathalie
APPLICANT: Ouellet, Catherine
TITLE OF INVENTION: Streptococcus Antigens
FILE REFERENCE: 055190-0044
CURRENT APPLICATION NUMBER: US/09/884,465A
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 60/212,683
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 384
SOFTWARE: PatentIn version 3.1
SEQ ID NO 373
LENGTH: 906
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Unknown Organism
US-09-884-465A-373

Query Match 7.5%; Score 130; DB 9; Length 906;
Best Local Similarity 23.6%; Pred. No. 0.083;
Matches 95; Conservative 47; Mismatches 152; Indels 108; Gaps 23;
1 MVAADAPAGCTDITNQL-TNVTVGIDSGTIVYPHOAG--YVKLNYGFSVPNSA----- 50
279 IASKDPEVSYDGTFTVPTSLAYKMASQTIYFPFHAGDTYLRVNPQFAVPKGTDALVRVF 338
51 -----VKGDTFKITVPKELNGLNGVTSTA--KVPPIAGDQVLANGVIDSDGNV 96
339 DEFGHAYLENNYKVGKILPIPK-LN-QGTRTAGNKIPVTF-----MANAYLDNOSTY 391
97 IYTF-----TD-----YVNTKDDVKATLTMPA-YIDPENVKKTGNVTLA 134
392 IVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPTSEKVEKEKLEKSETGST-- 449
135 TGIGSTANTKTVLVD--YEYKGF---YNLSIKGTIDQIDKTNNTYQTIYVNPSPGDNV- 188
450 ---SNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYL-----PSGEVIK 500
189 -----IAPVLTGNLKPNTD---SNALIDQONTSIKVKVDNAADLSSESYFVNPNFE 237
501 KNAADFTGEAPOGNGENKPSKGVSTGTVENQPTENK---PADSLPEAPNEKPVKPE--- 555
238 DVTNSVNIPTPNPNQYKVFENTPDDQITTPYIVVNGHIDP-----NSKGDALALRST 289
556 ---NSTDNGMLNP-----EGNVGSDPMLDPALEAPA-VDPVQEKLEKFTASYGLGLDSV 606
290 LYGNSNIWIRSWDNEVAFNNGSGSGDGDIDRPVPEQDPE 331

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:31 ; Search time 22.3014 seconds

(without alignments)
1426.837 Million cell updates/sec

Title: US-10-056-052A-4

Perfect score: 1732

Sequence: 1 MVAADAPAAAGTDITNQLTNV.....NGSGSGDIDKPVVPEQPDE 331

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

1 number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1727	99.7	933	2	S41539
2	1517	87.6	989	2	D89852
3	416.5	24.0	961	2	G90053
4	409	23.6	940	2	S19702
5	349	20.2	1092	2	T30214
6	331	19.1	1018	2	A32192
7	324.5	18.7	1038	2	H90053
8	323.5	18.7	1166	2	T28680
9	320.5	18.5	1141	2	E89824
10	312.5	18.0	877	2	F90070
11	257.5	14.9	1315	2	T28679
12	240.5	13.9	1385	2	D89824
13	229.5	13.3	953	2	C89824
14	182	10.5	1039	2	T30856
15	161.5	9.3	463	2	AG1542
16	157.5	9.1	462	2	AH1184
17	147.5	8.5	586	2	AD1458
18	141	8.1	1386	2	AC1533
19	137	7.9	439	2	AE1251
20	137	7.9	4152	2	T31102
21	135	7.8	1583	2	G85643
22	132.5	7.7	1457	2	D81019
23	132	7.6	1449	2	B81963
24	129	7.4	762	2	A34355
25	128	7.4	1158	2	AF1852
26	128	7.4	2021	2	A97859
27	127	7.3	542	2	D81401
28	127	7.3	793	2	AH1094
29	126.5	7.3	571	2	AH1094

outer membrane pro
hypothetical prote
hypothetical prote
probable cell surf
conserved hypothet
hypothetical prote
peptidoglycan boun
190K surface antig
hypothetical prote
probable adhesin h
hypothetical prote
extracellular neut
hypothetical prote
hypothetical prote
surface protein pr
extracellular prot
hypothetical prote
hypothetical prote
hypothetical prote
5'-nucleotidase, p
polymorphic oute
polymorphic membra
hypothetical prote
5'-nucleotidase, p
adhesin homolog lm
major ring-forming
outer membrane pro
hypothetical prote
probably celluloso
peptidoglycan anch
toxin-like outer m
hypothetical prote
hypothetical prote
surface layer prot
yfaL protein - Esc
lactococpin (EC 3.4
lactococpin (EC 3.4
lactococpin (EC 3.4
hypothetical prote
adhesin AIDA-I pre
IgA-specific serin
flagellar hook-ass
hemagglutinin/hemo
hypothetical prote
metalloproteinase
outer membrane ass
lactococpin (EC 3.4
FmtB protein (limpo
probable flagellin
probable structura
hypothetical prote
large repetitive p
hypothetical prote
S-layer protein -
hypothetical prote
secreted protein c
surface-layer 125K
hypothetical prote
crystalline surfac
probable cell surf
tail-host specific
hypothetical prote
hypothetical prote
metalloproteinase
alpha-glucosidase
probable invasin (p
probable invasin 2
probable peptidogl
cell surface antigen

ALIGNMENTS

RESULT 1
S41539
fibrinogen-binding protein - Staphylococcus aureus
N:Alternate names: clumping factor
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
C:Accession: S41539; S36630
R:McDevitt, D.; Francois, P.; Vaudeaux, P.; Foster, T.J.
Mol. Microbiol. 11, 237-248, 1994
A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus
A:Reference number: S41539; MUID:94224142; PMID:8170386
A:Accession: S41539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-933 <MCD>
A:Cross-references: EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g397526

Query Match 99.7%; Score 1727; DB 2; Length 933;
Best Local Similarity 100.0%; Pred. No. 3.3e-100; Mismatches 0; Indels 0; Gaps 0;
Matches 330; Conservative 0

QY 2 VAADAPAAAGTDTITNLTNTVTGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDFTFKITVP 61
DB 221 VAADAPAAAGTDTITNLTNTVKYTIIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDFTFKITVP 280

QY 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLMPAYID 121
DB 281 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLMPAYID 340

QY 122 PENVKKTGNVTLATIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTNNTYRQTIYV 181
DB 341 PENVKKTGNVTLATIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTNNTYRQTIYV 400

QY 182 NPSGDNVIAPVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSESYFVNPENFEDVTN 241
DB 401 NPSGDNVVLPAITGNLIPNTKSNALIDAKNTDQIKYRVNDNADLSESYFVNPENFEDVTN 460

QY 242 SVNTTFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNNSIIWRS 301
DB 461 QVRISFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNNSIIWRS 520

QY 302 MSWDNEVAFNNGSGGDGIDKPVVPEQDPE 331
DB 521 MSWDNEVAFNNGSGGDGIDKPVVPEQDPE 550

RESULT 2
D89852
fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (strain
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D89852
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D89852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-961 <KUR>
A:Cross-references: GB:BA000018; PID:gl3702452; PIDN:BA43593.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnbB

Query Match 87.6%; Score 1517; DB 2; Length 989;
Best Local Similarity 87.3%; Pred. No. 4.8e-87; Mismatches 25; Indels 0; Gaps 0;
Matches 288; Conservative 17

QY 2 VAADAPAAAGTDTITNLTNTVTGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDFTFK 57
DB 158 VAADAPAAAGTDTITNLTNTVTGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDFTFK 217

QY 58 ITVPEKLNGLNGVTSTAKVPPIMAG--DOVLANGVIDSDGNVIYFTDYVNTKDDVKATLT 115
DB 218 FTLSNDNVETHGISTLRKPEIKSSTEDKVMANGQVINERTYFTDYVNTKDDVKATLT 277

QY 116 MPAYIDPNVKKTNVTLATIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTNNTYRQTIYV 170
DB 278 LNLFDIPPTVTGKSGQVEVTGKNSKVEFDIKYLDGVKDRMG---VTVNGRIDTLNKK 333

QY 171 TNNTYRQTIYVNPSCDNVIAPVLTGNLKPNTDSNALIDQNTSIKVKYVDNAADLSESYF 230
DB 334 EEGKSFHAYFVPPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNNSIIWRS 389

QY 231 V---NPENFEDVTNSVNITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNKSGDLALR 287
DB 390 AKLDDTSKFDVTEKVNLSYTSNGGYTLNLGLDLN--SKDYVIKYEYEDQNAK--DLNFR 446

QY 288 STLGYNSNIWIR-----SMWDNEVAFNNGSGGDGIDKPVVPEQDPE 326
DB 521 MSWDNEVAFNNGSGGDGIDKPVVPEQDPE 550

Db 447 THLSGYHKYYPYYPYVQLTWNNGVAFYSNNAKGDCKPNDP 493

RESULT 4

S19702

fibronectin-binding protein B - Staphylococcus aureus

C:Species: Staphylococcus aureus

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999

C:Accession: S19702

R:Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.

Eur. J. Biochem. 202, 1041-1048, 1991

A:Title: Two different genes encode fibronectin binding proteins in Staphylococcus aureus

A:Reference number: S19702; MUID:92111475; PMID:1837266

A:Accession: S19702

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-940 <JOE>

A:Cross-references: EMBL:X62992; NID:949040; PIDN:CAA44726.1; PID:9581562

C:Keywords: fibronectin binding

Query Match 23.6%; Score 409; DB 2; Length 940;

Best Local Similarity 28.9%; Pred. No. 7.4e-18;

Matches 101; Conservative 72; Mismatches 131; Indels 46; Gaps 11;

Qy 2 VAADAPAGTDIRNLQNTVTVGIDSGT-----TVYPHQAGYVKLYNGFSPNSAV 51

Db 155 VVEETKATGTDVTK-----VEVEEGSEIVGHKQDTNWNPNHNAERVTLLKWKFGEGIK 209

Qy 52 KGDTFKITVPKELNLTSTAKVPPIMAGD-QVLANGVIDSGNVIYTFDVTNTKDDV 110

Db 210 AGDVFETLSNVETHGISLRLKVPKISTGQVMATGEIGERKRVYTFEYVQEKDL 269

Qy 111 KATLTPPAYIDPENVKKTGNVTATGIGSTTANKTVLVVDY-----EKYGFYKLSIKGTI 165

Db 270 TAELSLNLFDPTVTQKGNQVEVKLGTEVSKFNIQYLGVRDNGW---VTANGRI 325

Qy 166 DQDKTNTNTVQTIYVNPNGDNVAPVLTGNL-----KPNTDSNALIDQNTSIKVKYVDN 221

Db 326 DTLNKVDGKFSHEAFMKPNQNSLSSVTGQVTKGNKPGVN-----NPTVKVYKHIG 377

Qy 222 AADLSESYFVNPN---FEDVTSNVTIFPNQYKVEFTPDQITTPYIVVNGHIDP 278

Db 378 SDDLAESVYAKLDDVSFEDVTDMSLDFTNGSYSLNFDQ---SKNVYIKYEGYDS 435

Qy 279 NSKGDLALRSTLYGYNISIIWRSMWDNEVAFNNGSGGIDK---PVV 325

Db 436 NA-SNLEFQTHLGYNYNYTSNLTWKNVAFYSNNAQGDCKLKEPII 484

RESULT 5

T30214

fibronectin-binding protein - Staphylococcus epidermidis

C:Species: Staphylococcus epidermidis

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000

C:Accession: T30214

R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pel, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66, 2666-2673, 1998

A:Title: A fibrinogen-binding protein of Staphylococcus epidermidis.

A:Reference number: 220781; MUID:98261511; PMID:9596732

A:Accession: T30214

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1092 <NIL>

A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 20.2%; Score 349; DB 2; Length 1092;

Best Local Similarity 30.3%; Pred. No. 5.1e-14;

Matches 100; Conservative 61; Mismatches 139; Indels 30; Gaps 14;

Qy 13 ITNLTNTVTCI-DSGTTVYPHQAGYVKLYNGFSPNSAVKGDFTFKITVPKELNLTGVS 71

Db 287 VTDQ--SITEGYDDSEGVIRKAHDAENLIYDVTFEVDQVKVSGDPTMTVDIKNTVPISLTD 344

Qy 72 TAKVPPIMAGD-QVLANGVIDS-DGNVIYTFDVTNTKDDVKATLTMTPAYIDPENVKKTG 129

Db 345 SETPIKIDNSGEIATGTVDNKNKQITVFTDVKYENIKAHKLKLTSDKSKVPNNN 404

Qy 130 ---NVTLATGIGSTTANKTVLVVDYKGYKFNLSIKGTIDQIDKTNNTYRQTIYVNSGD 186

Db 405 TKLDVEYKTALSS--VNKTITVEYQRPNEARTANLQSMFTNIDTKNHTVEQTIYINPL-- 460

Qy 187 NVIAPVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSSEYFV-NPENFEDVTNSVNI 245

Db 461 RYSAKETNWNISNGDEGSIIDSTTIKVKYKVDNQNLDPDSNRIYDYSEYEDVTNDYA 520

Qy 246 TFPNPQYKVEFTNPDDQITTPYIVVNGHIDPNKSGD-----LALRSTLYGYNISII 298

Db 521 QLGNNDVNINFG---NIDSPYLIKISKYDPN-KDYTTIQOTVTMQTTINEYTGE-- 573

Qy 299 WRSMWDNEVAFNNGSGGIDKPVVPEQ 328

Db 574 FRTASYDNTIAFTSSGGQGG-DLP--PEK 600

RESULT 6

A32192

fibronectin-binding protein - Staphylococcus aureus

C:Species: Staphylococcus aureus

C:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 16-Feb-1997

C:Accession: A32192

R:Signaes, C.; Raucel, G.; Joensson, K.; Lindgren, P.E.; Anantharamaiah, G.M.; Hoeoek

Proc. Natl. Acad. Sci. U.S.A. 86, 699-703, 1989

A:Title: Nucleotide sequence of the gene for a fibronectin-binding protein from Staph

A:Reference number: A32192; MUID:8908998; PMID:2521391

A:Accession: A32192

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1018 <SIG>

A:Cross-references: GB:J04151

C:Keywords: fibronectin binding

Query Match 19.1%; Score 331; DB 2; Length 1018;

Best Local Similarity 24.6%; Pred. No. 6.1e-13;

Matches 82; Conservative 77; Mismatches 146; Indels 28; Gaps 11;

Qy 10 GTDITNQLTNTVTVGIDSG---TVYPHQAGYVKLYNGFSPNSAVKGDFTFKITVPKELN 65

Db 194 GTDVTSKVT-VEIGSIEGHNTNKVEPHAGRAVLKYLKLFENGHQGVDFDTLSNNVN 252

Qy 66 LNGVTSTAKVPPIMAGDQVLANGVIDSGNVIYTFDVTNTKDDVKATLTMTPAYIDPENV 125

Db 253 THGVSTARKVPETKNGSVVMTATGEVLEGGKIRYFTTNDIEDKVDVTAELINLFDPKTV 312

Qy 126 KKTGNVTLATGIGSTTANKTVLVVDY-EKYGFYKLSIKGTIDQIDKTNNTYRQTIYVNSP 184

Db 313 QTNGNQITITSLNEEQTSKELDVKYKDGIGNY-ANLNGSIETETKNKANNRSHVAFKPN 371

Qy 185 GDNVIAVLGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSSEYFV---PENFEDVT 240

Db 372 NGKTTSTVTGTGLMKGSNQNG---NQPKVRIFFYLGNNEDIKSVANTDTSKPKREV 427

Qy 241 NSV--NITFPNPQYKVEFTNPDDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNISII 298

Db 428 SNMSGNLNLQNGSYSLNI-----ENLQKTYVHVVDGEY-LNGTDEVDFTQMVGHPOLY 482

Qy 299 -----WRSMWDNEVAFNNGSGGIDKPVV 325

Db 483 KYYYDRGYTLTWGNGLVLYSNKANGNEKNGPII 515

RESULT 7

H90053

hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: H90053

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, R.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: E89824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1038 <KUR>
A:Cross-references: GB:BA000018; PID:gl3702453; PIDN:BA843594.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb

Query Match 18.7%; Score 324.5; DB 2; Length 1038;
Best Local Similarity 26.9%; Pred. No. 1.6e-12;
Matches 91; Conservative 68; Mismatches 154; Indels 25; Gaps 11;
10 GTDITNLTNTVGTIDS--GTTVPHQAGVVKLYNGFSPNSAVKGDFTKITVPKELNLN 67
189 GTDVTSKVTESGIEAPQGNKVEPHAGQVWLKYLKLFADGLKRGDYDFDTLSNNVTY 248
Qy 68 GVTSTAKVPPIMAGDQVLANGVDSGNIYFTDYVNTKDDVKATLTMPIAYIDPENVK 127
Db 249 GVSTARKVPEKNGSVYMATGEILGNIRYFTETNEIEHKVEVTANLEINLFIDPKTVQS 308
Qy 128 TGNVTLATGISTTANKTVLVDVEKYEKYNLSFKGIDIDDKTNTNTYRTIYVNPSCDN 187
Db 309 NGEOKITSKNGEETETIPVYVNGVSNYNTVNGSIETFNKESKFTTHAIYIKPMNGN 368
Qy 188 VIAPV-LTGNLKPNTSDNALIDQNTSIKVK-VDNAADLSSEYFVNPN--FEDVTNS 242
Db 369 QSNVSVTGTL---TEGSLNLAGGQPT-VKVEYLGKKDELPSQVYANTSDTNKEKVTKE 424
Qy 243 VN--ITFPNPNQKVENTPDDQITTYIVVNGHIDPNKSGDLALRSTLYG-----YNS 295
Db 425 MNGKLSVQDNGSYSLN---DKLDTYVIHYTGEYLGQS--DQVNFRTELYGPARYKS 479
Qy 296 NIWRS--MSWDNEAFNNGSGDGIDKVPVPEQDE 331
Db 480 YVYGGYELTWDNGLVLYSKADGNGKNGQIIONDFE 517

RESULT 8
T28680
Query Match 18.7%; Score 323.5; DB 2; Length 1166;
Best Local Similarity 31.2%; Pred. No. 2.2e-12;
Matches 111; Conservative 53; Mismatches 137; Indels 55; Gaps 18;
C:Species: *Staphylococcus aureus*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
A:Accession: T28680
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1166 <JOS>
A:Cross-references: EMBL:AJ005647; NID:el318793; PID:el318794; PIDN:CAA06652.1
C:Genetics:
A:Gene: sdrE

Query Match 18.7%; Score 323.5; DB 2; Length 1166;
Best Local Similarity 31.2%; Pred. No. 2.2e-12;
Matches 111; Conservative 53; Mismatches 137; Indels 55; Gaps 18;
2 VAADAPAAAGTDITNLTNTV-----VGIDSGTIVPHQAGVVKLYNGFSPNSAVKGDFT 56
Db 269 VAQPAAVASNNV--NDLIKVTQTKIKYGDGKDNVAAAHDKDIEYDTEFTIDNKVKGGDTM 327
Qy 57 KI-----TVPKELNLNGVTSTAKVPPIMAGD---QVLANGVID--SDGNVIYFTDYVNTK 107
Db 328 TINYDKNVIPSDL-----TDKNPDIDITDPSGEVIAKGTDFKATKQITTYFTDYVDKY 380

Qy 108 DDVKATLTMPAYIDPENV--KKTGNVTLATGISTTANKTVLVDVEKYEKYNLSIKCTI 165
Db 381 EDIKSRUTLTSYIDDKTVPNETSLNLTATAGKETSQN--VTVDYQDPMVHGDNSIOSIF 438
Qy 166 DQIDKTNNTYRQTIYVNP---SGDNVIAPVL-----TGNLKPNTSDNALIDQNTSIKV 216
Db 439 TKLDEKQTIQEQIYVNPPLKKSATNTKVDIAGSQVDYGNLKLNGS--TIID-QNTEK 496
Qy 217 KYVDNAADLSSEYFV--NPENFEDVTNSVNIITFPNPNQYKVEFNTPD--DQITTPYIVVNGH 274
Db 497 KYVNSDQQLPQSNRIYDFSQVEDVTSQ----FDNKKSFNNVATLDFGINSAYIIKVS 552
Qy 275 HIDPNSKGLD-----ALRST-LYGYNSNIWRSWSDNEAFNNGSGSGDGIDKP 323
Db 553 KYTPTSOGELDIAOGTSMRTTDKYG-----YNYAGYSNFIYVTSNDTGGGDTVKP 603
RESULT 9
E89824
hypothetical protein sdrE [imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
A:Accession: E89824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1141 <KUR>
A:Cross-references: GB:BA000018; PID:gl3700455; PIDN:BA841752.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: sdrE

Query Match 18.5%; Score 320.5; DB 2; Length 1141;
Best Local Similarity 30.4%; Pred. No. 3.3e-12;
Matches 108; Conservative 55; Mismatches 139; Indels 53; Gaps 17;
2 VAADAPAAAGTDITNLTNTV-----NVTVGIDSGTIVPHQAGVVKLYNGFSPNSAVKGDTEK 57
Db 264 VAQPAAVASNNVNDLITVTQTKIKYGDGKDNVAAAHDKDIEYDTEFTIDNKVKGGDTMT 323
Qy 58 I-----TVPKELNLNGVTSTAKVPPIMAGD---QVLANGVID--SDGNVIYFTDYVNTKD 108
Db 324 INYDKNVIPSDL-----TDKNPDIDITDPSGEVIAKGTDFKATKQITTYFTDYVDKYE 376
Qy 109 DVKATLTMPAYIDPENV--KKTGNVTLATGISTTANKTVLVDVEKYEKYNLSIKGTID 166
Db 377 DIKARLTLSYIDKQAVPNETSLNLTATAGKETSQN--VSDYQDPMVHGDNSIOSIFT 434
Qy 167 QIDKTNNTYRQTIYVNP---SGDNVIAPVL-----TGNLKPNTSDNALIDQNTSIKV 217
Db 435 KLDENKQTIQEQIYVNPPLKKSATNTKVDIAGSQVDYGNLKLNGS--TIID-QNTEK 492
Qy 218 KYVDNAADLSSEYFV--NPENFEDVTNSVNIITFPNPNQYKVEFNTPD--DQITTPYIVVNGH 275
Db 493 KVPNQLPQSNRIYDFSQVEDVTSQ----FDNKKSFNNVATLDFGINSAYIIKVS 548
Qy 276 IDPNSKGLD-----ALRST-LYGYNSNIWRSWSDNEAFNNGSGSGDGIDKP 323
Db 549 YTPTSOGELDIAOGTSMRTTDKYG-----YNYAGYSNFIYVTSNDTGGGDTVKP 598

RESULT 10
F90070
Clumping factor B [imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: F90070
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90070
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <KUR>
A:Cross-references: GB:BA000018; PID:g13702588; PIDN:BAB43728.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: clfB

Query Match 18.0%; Score 312.5; DB 2; Length 877;
Best Local Similarity 24.9%; Pred. No. 7.2e-12;
Matches 92; Conservative 69; Mismatches 132; Indels 77; Gaps 15;

6 APAAAGTDTITNLTNTVGTIDSGTTVYPHQAGYVKLNLYGFSVPNSAVRGDTFKITVPKELN 65
| | | | | : : : : : | | | | : : : : : | | | : : : : : | | : : : : : |
208 ADAAGTNVNDKVTASNEKLEK-TTFDPNQSGNTFMAANFTVDKVKSGDYF----- 257

66 LNVSTAKVPPIMAGDOVLANGVID-----SDGNVI----- 97
| | | | | : : : : : | | | : : : : : | | | : : : : : | | : : : : : |
258 -----TAKLP-----DSLTCNGDGYDSNSNTMPIADIKSTNGDVVAKATYDILTKTYT 306

98 YTFDYNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVLDVYEK----Y 153
: | | | | | | | : : : : : | | | : : : : : | | | : : : : : | | : : : : : |
307 FVTFDYNNKENINGQSLPFTDRAPKSGTYDANINADENFNKITYVSSPIAGI 366

154 GKFNLSIKGTIDQIDKTN--NTYRQTIYVNS-----GDNVIAPLVLTGNLKPNTDSNALI 207
: | : | : | : | : | : | | | | : : : : : | : | : | : | : | : | : | : | : | : |
367 DKPNGANISSQIIGVDTSAGONTYQYFVNPQKORVLGNTWV--YIKGYQDKIESSGKV 424

208 DQONTSKVKYVDNAADLSSEYFNP--ENFEDVNSV--NITFPNPQYKVFENTPDDQ 263
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
425 SATDTKLRIEFVNTSKLSDSYADPNDSNLKEVDFKNIYYEHPNVASIKFG-----D 480

264 ITTPYIVVANGHIDPNKSGDLALSTLYGVN----SNLIWRMSMDNEVAFNNGSGSGDG 319
| | | | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
481 ITKYVVLVEGHYDNTGKN---LATQVIOENVDPVTNRDYSIFGWNENNVYRGGSGADG 537

320 IDKPVVPEQP 329
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
538 -DSAVNPCKP 546

LT 11
T28679
fibrinogen-binding protein homolog - *Staphylococcus aureus*
C:Species: *Staphylococcus aureus*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: T28679
R:Josefsson, E.; McCrea, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of *S. aureus*.
A:Reference number: 220510; MUID:95098700; PMID:9884231
A:Accession: T28679
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1315 <JOS>
A:Cross-references: EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDN:CAA06651.1
C:Genetics:
A:Gene: sdrB

Query Match 14.9%; Score 257.5; DB 2; Length 1315;
Best Local Similarity 26.9%; Pred. No. 3.4e-08;
Matches 88; Conservative 51; Mismatches 157; Indels 31; Gaps 12;

13 ITNOLTNVTGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVRGDTFKITVPKELNNGVT-- 70
| | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Db 251 ITSNTTLTVVDADKNNKIVPAQ-DYLSLKSOITVDDKVKSGDYFTIKYSDTVQVYGLNPE 309
Qy 71 -----STAKVPPIMAGDOVLANGVIDSDGNVI-YTFDYNTKDDVKATLTMPAYIDPEN 124
| | | : | | : | | | | | | : | : | : | | | : | : | : | : | : | : | : |
Db 310 DIKNIGDIKDP---NNGETATAKHDTANNLITFTFDYDRFNSVQMGINYSYMDADT 366
Qy 125 VKKTGN-VTLATGIGSTTANKTTLVLDVYEKGKFNLSIKG----TIDQIDKTNNT--YRQ 177
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 367 IPVSKNDVERFNVITGNTTTKTANIQYPDYVNVNEKNSIGSAFTETVSHVGNKNGPYVKQ 426
Qy 178 TIYVNPSCDGVIAPLVLTGNLKPNTDSNAL--IDQONTSIKVKYVDNAADLSSEYFNPEN 235
| | | | | : : : : : | : : : : | | | : | : | : | : | : | : | : | : |
Db 427 TIYVNPSENSTNAKLQYAHSSYPNNIGINKDVTDIKIYQYKGYTLNKGKGYDVTNKE 486
Qy 236 FEDVNTSV--NITFPNPQYKVFENTPDDQITTPYIVVANGHID-PNSKGDALALRSTLYG 292
| | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 487 LTDVTNOYLQKITGYDNNNSAVIDFGNAD---SAYVVMVNTKFOYTNSSEPTLVQMATLS 542
Qy 293 YNSNIWRMSMDNEVAFNNGSGSGDG 319
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 543 STGN---KSVSTGNALGFTNNQSGGAG 566

RESULT 12

D89824
hypothetical protein sdrB [imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D89824
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D89824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1385 <KUR>
A:Cross-references: GB:BA000018; PID:g13700454; PIDN:BAB41751.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: sdrB

Query Match 13.9%; Score 240.5; DB 2; Length 1385;
Best Local Similarity 26.1%; Pred. No. 4.2e-07;
Matches 86; Conservative 54; Mismatches 154; Indels 35; Gaps 14;

Qy 13 ITNOLTNVTGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVRGDTFKITVPKELNNGVT-- 70
| | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 251 ITSNTTLTVVDADKNSKTIVPAQ-DYLSLKSOITVDDKVKSGDYFTIKYSDTVQVYGLNPE 309

Qy 71 -----STAKVPPIMAGDOVLANGVIDSDGNVI-YTFDYNTKDDVKATLTMPAYIDPEN 124
| | | : | | : | | | | | | : | : | : | | | : | : | : | : | : | : | : |
Db 310 DIKNIGDIKDP---NNGETATAKHDTANNLITFTFDYDRFNSVQMGINYSYMDADT 366

Qy 125 V---KKTGNVTLATGIGSTTANKTTLVLDVYEKGKFNLSIKG----TIDQIDKTNNT--Y 175
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 367 IPVDKK--DVPFVSITIGNQITTTTADITYPAKADNNNSIGSAFTETVSHVGNVEDPGYY 424

Qy 176 RQTIYVNPSCDGVIAPLVLTGNLKPNTDSN-ALIDQONTSIKVKYVDNAADLSSEYFNP 233
| | | | | : : : : : | : : : : | | | : | : | : | : | : | : | : | : |
Db 425 NQVYVNPMDKLGAKLKEAYHPKYPTNIGNQINQNVNINIKIVRPEGYTLNKGKGYDVT 484

Qy 234 ENFEDVNTSV--NITFPNPQYKVFENTPDDQITTPYIVVANGHID-PNSKGDALALRSTL 290
: | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 485 NDLVDVTDDEFKNKMTYGSNOSVNLDFG----DITSAYVVMVNTKFOYTNSSEPTLVQMAT 540

Qy 291 YGYNSTIWRMSMDNEVAFNNGSGSGDG 319
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 541 LSSTGN---KSVSTGNALGFTNNQSGGAG 566

```

RESULT 13
C98824
hypothetical protein sdrC [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C98824
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cuf, L.; Oguc
ma, A.; Mizutani-Uti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A99758; MUID:21311952; PMID:11418146
A:Accession: C98824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-953 <KUR>
A:Cross-references: GB:BA000018; PID:gl3700453; PIDN:BA41750.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: sdrC

Query Match 13.3%; Score 229.5; DB 2; Length 953;
Best Local Similarity 24.9%; Pred. No. 1.2e-06;
Matches 84; Conservative 56; Mismatches 157; Indels 41; Gaps 12;

QY 6 APAAGTDTITNOL--TNVTVGIDSG-----TTVPHQAGYVKLNYGFSVPNSAVKGD 55
DB 178 APOQGTNVNDKVFHTNIDIAIDKGHVKTNTGTEFWATSSDVLKLNKANTIDDSVKEGDT 237

QY 56 FKITVPKELNLNGVTSTAKVPPIM--AGDQVLANGVIDSDGN--VIYFTDYVNTKDDVKAT 113
DB 238 FTEKYGVFRPGSVRLPSQTNLYNAOAGNIIAKGIYDSKTTNTTYFTNYVDQYTNVSGS 297

QY 114 LTPMAYIDPEN--YKKTG---NVILATIGSTTANKTVLVDEYKGYKFNLSIKGTDOI 168
DB 298 FEQAFARENATTDKTAAYKEVTL---GNDTYSKDVIYDYGKQGOQLISSTNYINNE 353

QY 169 DKTNTTYQTIVNPSGDNVIAPIVLTGNL---KPTDSNALIDQONTSIKVKYVDNAADL 225
DB 354 DLSRN---MTVYVNPQPKTYKTEFVNTLTGKFNPDAK-----NFKIYETDQNOF 402

QY 226 SESFVNPENFEDVTVNSNITFPNPNQ--YKVEFNTPDQITPTIYIVVNGHIDPNSKGL 284
DB 403 VDSEPTDTSKLKDTQGDVIVSYNDKNTATVDLLNGSSDKQXIIQOVAYPDNSSTDNG 462

QY 285 ALRSTLYCYNINIRSNWSWNEVAFNNGSGDGIDK 322
DB 463 KIDYTLTQNG-----KSSWSNSYSNVNGSGSTANGDQK 495

RESULT 14
T30856
protein F2 - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30856
R:Jaffe, J.; Natanson-Yaron, S.; Capaton, M.G.; Hanski, E.
Mol. Microbiol. 21, 373-384, 1996
A:Title: Protein F2, a novel fibronectin-binding protein from Streptococcus pyogenes, po
A:Reference number: 220907; MUID:97011581; PMID:8858591
A:Accession: T30856
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1039 <JAF>
A:Cross-references: EMBL:U31980; NID:gl654115; PID:gl654116; PIDN:AAC44522.1
C:Genetics:
A:Note: PRPF2

Query Match 10.5%; Score 182; DB 2; Length 1039;
Best Local Similarity 21.9%; Pred. No. 0.0013;
Matches 81; Conservative 78; Mismatches 149; Indels 62; Gaps 18;

QY 8 AAGTDITN--QLTNTVTGVI-----DSGTTVPHQAGYVKLNYGFSVPNSAV---K 52

```

```

Db 21 AGSKDVSSSLQLENPKMSVSVKGYGEOEKTSNSADFYRNHAAAYFKMSFELQKQKSETINP 80
QY 53 GDTKTIYVPKELNLNGVTSTAKVPPIM--AGDQVLANGVIDSDGN--VIYFTDYVNTKDD 109
DB 81 GDTFVLQDLRLNPKGISQ--DIPKIIYDSENSPLAIGKYDAKTHQLTFTYFTNYTAGDLK 138
QY 110 VKATLTMPAYIDPENVKKTGNVT--LATGIG--STTANKTVLVDEY---EKYCKFYN---L 159
DB 139 VOLSAAELSLFLENKEVLENTNISDFKSTIGGOEITYKGTNVNLYGNESKTESNYITNGLS 198
QY 160 SIKGTIDQIDKTNNTYROTIVYVNPNGDNVIAPIVLTGNL-----KPTDSNALIDQON 211
DB 199 NVGGSIESYNTETGEFVWYVYVNPRTNIPYAVL--NLWGAFAKTAQAGENDNSLVSSAQL 256
QY 212 TSIKVKYVDNAADLSSESYFVNPENFEDVTVNSNITFPNPN-----QYKVEFNTPDDQIT 265
DB 257 TGYDIYEVPHNYRLPTSYGVDISRL--NLKDKLEAKLPQGSTQGANKRLRIDFG--ENLQG 313
QY 266 TPYIVVNGHIDPNSKGLALRSTLYGYNSNIIMPS-----MSWDNEVAFNNGSGS 316
DB 314 KAFVVKVTGKADQSGK--ELIVOSHLSFPNN---WGSYKTLRPNHSHVSTNEIALUSPSKGS 369
QY 317 GDGIDKPVVP 326
DB 370 GSGTSEFTKP 379

RESULT 15
AG1542
wall associated protein precursor (LPXTG motif) homolog lin0879 [imported] - Listeria
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG1542
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1542
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-463 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96111.1; PID:gl6413329; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0879

Query Match 9.3%; Score 161.5; DB 2; Length 463;
Best Local Similarity 24.9%; Pred. No. 0.0081;
Matches 88; Conservative 48; Mismatches 129; Indels 89; Gaps 18;

QY 9 AGTDI--TNOLTNVTGVIDSGTIVYP--HOAGYVKLNYGFSVPNSAVKGDTEKITVPKELNL 66
DB 25 AATDYGSSFFTNVALQNGDPVTFNFKNSKRVAYDFVITQPVVSGGTMTLTIPDQKL 84

QY 67 NGVTSTAKVPPIMAGDQVLANGVID--SDGNVIYFTDYVNTKDDVKATLTMPAYIDPEN 125
DB 85 ---INFGGFPVNDASGNTIANATIDPATGTTLTFTTHVNTHTNLGSLFYNATFNASKNI 141
QY 126 KK-----TGNVTLATIGSTTANKTVLVDEYKGYKFNLSIKGTI 165
DB 142 QTDQVNPVQFPVNNNTTQINTYISKVTGGGTGPTI-----VFKQGRM 185

QY 166 DQIDKTNNTYROTIVYVNPNGDNVIAPI-----VLTGNLKNKPTDSNALIDQONTSIKVKYVDN 221
DB 186 D--DKDTSIHWTNLTN-----NALTPIDNAIYTDLT--GTQONLL---GNATIK--YRDAN 233
QY 222 AADLSSEYFVNP-----ENFEDVTVNSNITFPNPNQYKVEFNTPDDQITPTIYIVVNGH 275

```


Db 234 KVLSTN--VOPIALDSNRNFELTIGTLN-----NOSVV--IYDITKITTOKSYTN-- 281
Qy 276 IDPNSKGDALRLSTLYGNSNIWRSMWSWNEAFNNGSGSGDGIDKPVVPEOP 329
Db 282 -----KATLSDNLDVARNATVND--YSGSGGQTGTPTTPPPVKEEP 322

RESULT 16
AD1458
wall associated protein precursor (LPXTG motif) homolog lmo0880 [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AH1184
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1184
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC98958.1; PID:g16410283; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0880

Query Match 9.1%; Score 157.5; DB 2; Length 462;
Best Local Similarity 22.0%; Pred. No. 0.014;
Matches 80; Conservative 52; Mismatches 123; Indels 109; Gaps 16;

Qy 9 AGTDI-TNQLTNVTGIDSGTIVYVPHOA-----GYVKLVNGFSVPNSAVKGDFFKTVTP 61
Db 25 AADYGSSTFNVLQNG-----EQATNFKNSKVRVAYDFVITQPVASGETMTLITP 79
Qy 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDS-DGNVIYTFDYVNTKDDVKATLTMPAYI 120
Db 80 DQLKL---INVGFPFLMDSQGTIANATIDVGTITLTFTDYVNTHTDLSGLFYATFP 136
Qy 121 DPENVKGTNTLATIGSTTANKTVLVD-----YEKGRF-----YN 158
Db 137 NSKNIQTDQVNPVAFVKNNTQTVPIYSKNSGGTGTPTVFKQGRMDKDLISLHWT 196
Qy 159 LSIGTTIDQIKTNTVTROTIVYVNSGDNVIA-----PVLGNLKPNTDSNALI 207
Db 197 VTLNALTPID--NAVYTDIL--GSGQNLGSGATIKYRANKKVIATNIQP-----IALD 247
Qy 208 DQNTSIKVKVDNAADLSSEYFVNPFEDVTSNVNITFNPNOYKVEFNTDQDITTP 267
Db 248 ADRNFELSIGALNN-----QSVVITY-----DTKITK 275
Qy 268 YIVVNGHIDPNSKGDALRLSTLYGNSNIWRSMWSWNEAFNNGSGDGIDKPVVPE 327
Db 276 QKSYTN-----KATLSDNLDVSR-----NATVNDYSGG-GQGTGTPPAPP 316
Qy 328 QPDE 331
Db 317 VKEE 320

RESULT 17
AD1458
probable peptidoglycan bound protein (LPXTG motif) lin0203 [imported] - Listeria innocua
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD1458
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1458
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-586 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC95436.1; PID:g16412622; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0203

Query Match 8.5%; Score 147.5; DB 2; Length 586;
Best Local Similarity 25.4%; Pred. No. 0.084;
Matches 85; Conservative 36; Mismatches 117; Indels 97; Gaps 17;

Qy 36 GY---VVLNNGFSPNSA-VK-GDTFKITVPKELNLNGVTSTAKVPPIMAGDOVLANGVI 90
Db 62 GYDVSQVQHYTWAIPNSSNVKAGDTMQFVLPPELQI--VTDLDFSLDKDHKGN-VVGNVVA 118

Qy 91 DSD-GNVIYTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKTVLVD 149
Db 119 TKDTGKVVIITFDVEKNSDVSGLDFWSWMDKSLVEGNEKVPVLPVNGTGTITVEV- 177

Qy 150 YEKYKFNLSIKGTIDQIDKNTNTYRTQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQ 209
Db 178 -----GGKNOIDPTETLYKYG-WANAKNPDLIQWVVRVYAKENIQNAVYED 223

Qy 210 QNTSIKVKVDNAADLSSEYFVNPFEDVTSNVNITFNPNOYKVEFNTDQDITTPYI 269
Db 224 -----FVGPQKVIDF-NSIQAVHG-----BFD-PDDNFTP--- 251

Qy 270 VVNGHIDPNSKGDALRLSTLYGNSN-----IIWRSMWDN----- 306
Db 252 ----GAAPVSSD-----IIQTTDGFKNLGNLTDVSKISYITTTDNGASPSYTNKGKLTG 303

Qy 307 -----EVA--FNNGSGSGDGIDKPVVPEQDE 331
Db 304 DNVYVTQIEIVATPTSGSGSGGEGTGTGVELTKTDD 338

RESULT 18
AC1533
surface protein (LPXTG motif) [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AC1533
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1533
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1386 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96035.1; PID:g16413254; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0803

Query Match 8.1%; Score 141; DB 2; Length 1386;
Best Local Similarity 20.1%; Pred. No. 0.69;
Matches 84; Conservative 40; Mismatches 117; Indels 176; Gaps 14;

Qy 29 TVYPHOAGYKLVNGFS-VPNSAVKGDFTFKITVPKELNLNGVTSTAKVPPIMAGDOVLAN 87
Db 924 TIDTTKAGNDITYSYSGVTRSTELSTKITVTVTKK----NOVNLEAKDSTLYEGDKWIAK 979

Qy 88 ----GVIDSGN----- 95
Db 980 DNFSATDKDGTVDFAKAEIVKGTNTTKAGTYKITYSAGISKITITVTVLANQTKIVAK 1039
Qy 96 --VIYTFDYVNTKDDVKATITMPAYIDPENVKGTGNVTLAT-----GIGSTTA 142
Db 1040 DLTIYEGDNWKEQDNFVSATDKFGQIDFNSVKVTSVDIQTGPKYRITYSIEG-ASTTI 1098
Qy 143 NKTVLVDY-----EKYGFYNLS---IKGTIDQIDKTN 173
Db 1099 TVTLADOSNLVAKNSTIYVGDWKOSDNFVSATDKYKPKIDLSLLVTGTVD--TTTPG 1156
Qy 174 TYROTIVN-----PSGDNVAPV----- 192
Db 1157 EYEITYSVNGLTITVTVKENQANIVAEDSTIHAKESWKAADNFVSATDKTKNIGLSS 1216
Qy 193 --LFGNLKPNTDSN-----ALIDQONTSIKV---YKVDNAADL 225
Db 1217 VSVTGKVDINTPGNVEITYTIDGVSSTITVTVLVNHSQIEAHDAKIKVGTSMKPEDSFIL 1276
226 SESYFVNPFENFVNSNITFPNPNQYKVEFNTPDDOITPIYIVVNGHIDPNKSG 282
Db 1277 AKDFGETADFSVTVTEGTVDTTPGKYQIYITIDGVSVTITVIV-----EDNSKG 1327
RESULT 19
AE1251
probable peptidoglycan bound protein (LPXTG motif) lmol413 [Imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE1251
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1251
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-439 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99491.1; PID:g16410842; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmol413

Query Match 7.9%; Score 137; DB 2; Length 439;
Best Local Similarity 23.5%; Pred. No. 0.25;
Matches 81; Conservative 48; Mismatches 128; Indels 88; Gaps 20;

Qy 31 YPHOAGYVKL-NYGFSPNSAVKGD-TFKITVPKEL-----NLNGVTSTAKV--PPIMAGD 82
Db 63 YPGNAAYLFNNKYNITDYVDNFAFLKPLKPTLEAGDPLNFTITGNVLDPVAPGQ 122
Qy 83 Q--VLANGVIDSD--GNVIYTFDYVN---TKDDVKATL-----TMPAYIDPENVKVTGN 130
Db 123 ESKYLAGFPFPIKANTQVNVDETNOTLATSLSKLGETYTTSPKADIGYQVKET-- 180
Qy 131 VTLATGISTTANKTVLVDYKYGFYNLSIKGTIDQIDKTNNTYROTIIYVNPSCDNVIA 190
Db 181 PTNATGT-YTTNTETIQINY---VYEKTAVEGANVSVE-----YINEATNESIA 225
Qy 191 P--VLITGNL-----KPNVDSNALIDQONTSIKVYKVDNAADLSSEYF 230
Db 226 PTETLSKIGITFOAEVKEIDGYELSQVPSQSGTYDQTSQVIFKFKYKKTTPD----- 279
Qy 231 VNPENFEDVNSNITFPNPNQYKVEENTPDQITPIYIVVNGHI-----DPNSKGDAL 286
Db 280 -----EVAKDVTVT-----YK---DTKNQIADP--VILKGDIGSTFTEAKSPGY 321
Qy 287 RSTLYGYSNIIWRMSWDNEAFNNGSGGIDGDKPVVPQPD 331

Db 322 KLTKTPSNHSSVFTSDQSEVEVYSKDDAV---ITPPVTPVNPDK 363
RESULT 20
T31102
filamentous hemagglutinin 1 - Haemophilus ducreyi
C:Species: Haemophilus ducreyi
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T31102
R:Ward, C.K.; Lumley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.; J. Bacteriol. 180, 6013-6022, 1998
A:Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.
A:Reference number: 220984; MUID:99030326; PMID:9811662
A:Accession: T31102
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-4152 <WAR>
A:Cross-references: EMBL:AF057695; NID:g3929017; PID:g3929018; PIDN:AAC79757.1
C:Genetics:
A:Gene: lspA1

Query Match 7.9%; Score 137; DB 2; Length 4152;
Best Local Similarity 22.0%; Pred. No. 5.6;
Matches 72; Conservative 55; Mismatches 129; Indels 72; Gaps 15;

Qy 4 ADAPAAAGTDTITNQLTNTVTVGIDS-----GTTVYHQAGYVKLYNFGFSPNSAVKGDTPKI 58
Db 619 SDFTSNGSKLVDAQNLTNTVNNFNITQSGSIILH--GNVTLNAKGNFTNSG-----NL 670
Qy 59 TVPKELNLTNGVTSTAKVPPIMAGD--OVLANGVIDSDGNVIYTFDYVNTKDDVKATLTM 116
Db 671 TTMKELNISNTESINAGNLTTGNLEVHSNTTVKNDGKLVSIENLNISSKTDFTNNGTL 730
Qy 117 PAYIDPENVKGTNTVLTATGISTTANKTVLVDYKYGFYNLSIKGTIDQIDKTNNTYR 176
Db 731 LG-LEALKIASGGNFTNASN--GSLASNKSL----DIYGN--NFTNNGTIESVKSLNITNN 782
Qy 177 QTIYVNPSCDNVIAPIVLTGNLKPNTDSNALIDQONTSIKVYKVDNAADLSSEYFVNPEN- 235
Db 783 YT-FIN-----NATIKSYGLNIT--SOGNFTNDSNG 811
Qy 236 -----PEDVTNSVNITFPNPN-----QYKVEFNTPDDQITPIYIVV-VNGHIDPNS 280
Db 812 TVMSHDLNITISQANII--NKNLLAGGOGNLTAKNITNDSNSTAIATVLSNNDINLA 869
Qy 281 KGDIALRSTLYGYSNIIWRMSWDNEV 308
Db 870 NNVYNGEITYSQAGNISVEAKLLHNDV 897

RESULT 21

G86643
hypothetical protein ybeF [Imported] - Lactococcus lactis subsp. lactis (strain IL140)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86643
R:Bolet, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86643
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1983 <STO>
A:Cross-references: GB:AE005176; PID:g12723000; PIDN:AAK04249.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ybeF

Query Match 7.8%; Score 135; DB 2; Length 1983;
Best Local Similarity 18.7%; Pred. No. 2.7;
Matches 85; Conservative 50; Mismatches 118; Indels 202; Gaps 17;

A34355
cell surface protein precursor - Acetogenium kivui
C:Species: Acetogenium kivui
C:Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 08-Oct-1999
C:Accession: A34355
R:Peterson, J.; Peters, M.; Lottspeich, F.; Baumeister, W.
J. Bacteriol. 171, 6307-6315, 1989
A:Title: S-layer protein gene of Acetogenium kivui: cloning and expression in Escherichia coli
A:Reference number: A34355; MUID:90036724; PMID:2681162
A:Accession: A34355
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-762 <PEP>
A:CROSS-references: GB:M31069; NID:g141844; PIDN:AAA21930.1; PID:g141845
C:Superfamily: S-layer repeat homology
F:32-85/Domain: S-layer repeat homology <SLR1>
F:96-150/Domain: S-layer repeat homology <SLR2>

Query Match 7.4%; Score 129; DB 2; Length 762;
Best Local Similarity 21.6%; Pred. No. 1.7;
Matches 79; Conservative 54; Mismatches 113; Indels 120; Gaps 19;

Qy 15 NQLTNVTGIDS-GTIVYVPHOAGYVKNLYGFSVPNSAVKGTFTKIVPKELN-----LNG- 68
Db 280 NDVVSFTGQDSVGTIVY-----KNDNKTAKVDDNAYVLYNGY 319
Qy 69 VTSTAKVP-----PIMAGQVLANGVIDSDG---NVIYFTDYVN--TKDDVKATLTM 116
Db 320 LTKYSKVTVKECAEVTIINNYLVINGSYDNSTIVYNDVQSGDKYLNKDSNYELKGTIVT 379
Qy 117 PAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKFKYNLS-----IK 162
Db 380 -----TGAVSKVTDI---KAN-----DYIYCKQYDVNGVYGTIVYVVRNQVT 420
Qy 163 GTIDIDKNTNRYRTIYVNSGDNVIAPVLGN-----LKPNTDSNALIDQOOTSIVKYK 218
Db 421 GTVTEKSVSGTYKASI-----DNVSYTVADNNVWNLQEPGKVTIVLNKDNVIVGI-- 472
Qy 219 VDNAADLSSEYFVNPFENFEDVTNSVNIITFPNPQYKVEFTPD-----DQI 264
Db 473 ----SSTTTTAVNYAIFKESDPTAWFA-----KVLILPDAEKVFDVAVSDYDKV 523
Qy 265 TTPYIVVYNGHIDPNSK-----GDLALSTLYGYNLSIKGTIDQIDKTNNTYR 313
Db 524 NLAEGTIVTYTVDANGKLDIQANDQPFSSASYKADAKVLTEGSTTYIITDNTVLLNNT 583
Qy 314 SGSGDG 319
584 S---DG 586

RESULT 25
AF1852
hypothetical protein alr0367 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AF1852
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF1852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1158 <KUR>
A:CROSS-references: GB:BA000019; PIDN:BAB72325.1; PID:g17129712; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr0367

Query Match 7.4%; Score 128; DB 2; Length 1158;

Best Local Similarity 22.0%; Pred. No. 3.5;
Matches 76; Conservative 42; Mismatches 123; Indels 104; Gaps 17;

Qy 2 VAADAPAAAGTDITNQLTNVTV---GDSGTTVYVPHOAGYVKNLYGFSVPNSAVKGTFTK 57
Db 713 IGAGAVNGGDDIDITQORLTLTNGGGID--TWVSPPEG-----NLPGGQGGGTIR 761
Qy 58 ITVPKELNLNGVTSTAKVPPIMAGQVLANGVIDSDGNVIVTFTDYVNTKDDVK-ATLTM 116
Db 762 INATDAVTISGTNANGFESALITETQ---GGAFGSGDIIIN-TDYFRVVDNGSVSTLT- 816
Qy 117 PAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKFKYNLSIKGTIDQIDKTNNTYR 176
Db 817 -----GNSSDAGNLIINTRVNEVLNGAAVVTGTESICKAGDITINAT-DKI----- 861
Qy 177 QTIYVNPSSGDNVIAPVLGTGNLKPNTDSNALIDQOOTSIVKYKVDNAADLSSEYFVNPF 236
Db 862 -----NLSNTDLHG-----NTGL-LASTTSSGDAGNIYL----- 890
Qy 237 EDVTNSVNIITFPNPQYKVEFTPDQITTPVIV-----LPAPVLVATLSQSGGIAGNINIVATGNVANN- 935
Db 891 --FTTELNLAPNSN-----NIWRMSWDNEVAFNNGSGSGDGI 320
Qy 282 GDLALSTLYGYNLS-----GDLALSTLYGYNLSIKGTIDQIDKTNNTYR 320
Db 936 GLISARSEQAGGGNLSVARNINLRNNS-DIRTDLFSGSGKGGI 979

Search completed: June 23, 2003, 14:04:13
Job time : 27.3014 secs


```

FNBA_STAAR          STANDARD;          PRT; 1018 AA.
ID  FNBA_STAAR          PRT; 1018 AA.
AC  P14738;
DT  01-APR-1990 (Rel. 14, Created)
DT  01-APR-1990 (Rel. 14, Last sequence update)
DE  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Fibronectin-binding protein precursor (FNBP).
GN  FNBA.
OS  Staphylococcus aureus.
OC  Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX  NCBI_TaxID=1280;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=89098998; PubMed=2521391;
RA  Signaes C., Raucel G., Joansson K., Lindgren P.-E.,
RA  Anantharamiah G.M., Hoeck M., Lindberg M.;
RT  "Nucleotide sequence of the gene for a fibronectin-binding protein
RT  from Staphylococcus aureus: use of this peptide sequence in the
RT  synthesis of biologically active peptides."
RT  Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).
CC  -!- FUNCTION: THE ABILITY OF BACTERIA TO BIND FIBRONECTIN HAS BEEN
CC  PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE
CC  WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO
CC  THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S.AUREUS,
CC  THUS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE
CC  INVASION.
CC  -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC  an amide bond (Potential).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: J04151; AAA26632.1;
CC  InterPro: IPR004237; Fn_bind.
CC  InterPro: IPR001899; Gram_pos_anchor.
CC  Pfam: PF00746; Gram_pos_anchor; 1.
CC  Pfam: PF02986; Fn_bind; 1.
CC  TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
CC  TIGRFAMs: TIGR01168; YSTRK_signal; 1.
CC  PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
CC  Cell wall; Peptidoglycan-anchor; Repeat; Signal.
CC  SIGNAL
CC  1 36
CC  CHAIN 37 985 FIBRONECTIN-BINDING PROTEIN.
CC  PROPEP 986 1018 REMOVED BY SORTASE (POTENTIAL).
CC  REPEAT 545 574 B-1.
CC  REPEAT 575 604 B-2.
CC  DOMAIN 745 878 4 X APPROXIMATE TANDEM REPEATS,
CC  FIBRONECTIN-BINDING DOMAIN.
CC  REPEAT 745 782 D-1.
CC  REPEAT 783 820 D-2.
CC  REPEAT 821 859 D-3.
CC  REPEAT 860 878 D-4. (INCOMPLETE).
CC  DOMAIN 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).
CC  REPEAT 879 892 WR1.
CC  REPEAT 893 906 WR2.
CC  REPEAT 907 920 WR3.
CC  REPEAT 921 934 WR4.
CC  REPEAT 935 948 WR5.
CC  SITE 982 986 LPXTG SORTING SIGNAL (POTENTIAL).
CC  MOD_RES 985 985 AMIDE-LINKED TO CELL WALL (POTENTIAL).
CC  SEQUENCE 1018 AA; 111780 MW; 58175E0020E81F1F CRC64;
CC  -----
Query Match 19.1%; Score 331; DB 1; Length 1018;
Best Local Similarity 24.6%; Pred No. 3.3e-13;
Matches 82; Conservative 77; Mismatches 146; Indels 28; Gaps 11;
QY  .10 GTDITNQLNTVTVGIDSG-----TTVYPHQAGYKLVNYGFSVPNSAVKGDFTKITYPKELN 65

```

```

Db 194 GTDVTSKYT-VEIGSIEGHNNTNKVEPHAGORAVLKYLKPEGLHOGDYDFDTLSNNVN 252
QY 66 LNVGTVSTAKVPPIMAGDQVLANGVYDSGNIYITFTDYVNTKDDVKATLTMPAYIDPENV 125
Db 253 THGVSTARKVPEIKNGSVVMATGEVLEGGKIRYTFETNDIEDKVDVTALEINLFDPKTV 312
QY 126 KKTGNVTLATGIGSTANKTVLVDY-EKYGRFYNLSIKGTIDQIDKTNNTYQTIYVNPVS 184
Db 313 QTNGNQTTITLNEEQTSKELDKYKDGIGNYI-ANLNGSIETFNKANNRFSHVAFIKPN 371
QY 185 GDNVIAPVLTGMLKPKNTSDNALIDQONTSIKVKY-VDNAADLSSEYFVN---PENFEDVT 240
Db 372 NGKTTSTVTGTLMKGSNQNG-----NQPKVRIEYVLGNNEIDAKSVANTTDTSAFKEVT 427
QY 241 NSV---NITFPNPQYKVEFNTPDDQITPTTYIVVNGHIDPNKSGDLALRSTLYGNSNII 298
Db 428 SNMSGNLNLQNGSYSLNI-----ENLDRKTYVVHYDGEY-LNGTDEVDFRTQMVGHPEQLY 482
QY 299 -----WRSMSSWDNEVAFNNGSGSGDGIDKPKV 325
Db 483 KYYDRGYTLTWNDGLVLYSNKANGNEKNGPII 515

RESULT 2
SLAP_ACEKI          STANDARD;          PRT; 762 AA.
ID  SLAP_ACEKI          PRT; 762 AA.
AC  P22258;
DT  01-AUG-1991 (Rel. 19, Created)
DT  01-AUG-1991 (Rel. 19, Last sequence update)
DE  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Cell surface protein precursor (S-layer protein).
OS  Acetogenium kivui.
OC  Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC  Thermoanaerobacteriaceae; Thermoanaerobacter.
OX  NCBI_TaxID=2325;
RN  [1]
RP  SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC  STRAIN=DSM 2030;
RX  MEDLINE=90036724; PubMed=2681162;
RA  Peters J., Peters M., Lottspeich F., Baumeister W.;
RT  "S-layer protein gene of Acetogenium kivui: cloning and expression in
RT  Escherichia coli and determination of the nucleotide sequence."
RL  J. Bacteriol. 171:6307-6315(1989).
RN  [2]
RP  PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
RC  STRAIN=DSM 2030;
RX  MEDLINE=92281680; PubMed=1596358;
RA  Peters J., Rudolf S., Oschkinat H., Mengele R., Sumper M.,
RA  Kellermann J., Lottspeich F., Baumeister W.;
RT  "Evidence for tyrosine-linked glycosaminoglycan in a bacterial
RT  surface protein."
RL  Biol. Chem. Hoppe-Seyler 373:171-176(1992).
RN  [3]
RP  DOMAINS.
RX  MEDLINE=94156823; PubMed=8113161;
RA  Lupas A., Engelhardt H., Peters J., Santarius U., Volker S.,
RA  Baumeister W.;
RT  "Domain structure of the Acetogenium kivui surface layer revealed by
RT  electron crystallography and sequence analysis."
RL  J. Bacteriol. 176:1224-1233(1994).
CC  -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC  OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC  -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC  S-LAYER WITH HEXAGONAL SYMMETRY.
CC  -!- PTM: THE CARBOHYDRATE CONTENT OF THIS PROTEIN IS ABOUT 8% WHICH
CC  CORRESPOND TO ABOUT 40 TO 50 SUGAR MOLECULES PER MONOMER. O-LINKED
CC  GLYCANS CONSIST OF GLC, GALNAC AND GLCNAC.
CC  -!- SIMILARITY: CONTAINS 3 S-LAYER HOMOLOGY (SLH) DOMAINS.
CC  -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its

```


RESULT 5
 OMPA_RICRI STANDARD; PRT: 2249 AA.
 ID OMPA_RICRI STANDARD; PRT: 2249 AA.
 AC P15921;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DE Outer membrane protein A precursor (190 kDa antigen) (Cell surface antigen) (rOmpA) (rOmpA).
 GN OMPA.
 OS Rickettsia rickettsii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=783;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R.
 RX MEDLINE=90354033; PubMed=2117568;
 RA Anderson B.E., McDonald G.A., Jones D.C., Regnery R.L.;
 RA "A protective protein antigen of Rickettsia rickettsii has tandemly repeated, near-identical sequences.";
 RA Infect. Immun. 58:2760-2769(1990).
 CC -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY.
 CC -!- SURCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-LAYER WITH HEXAGONAL SYMMETRY.
 CC -!- PTM: GLYCOSYLATED (PROBABLE).
 CC -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMP B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M31227; AAA26380.1;
 DR PIR; A41477; A41477.
 DR InterPro; IPR003858; rOmpA_rOmpB.
 DR Pfam; PF02708; rOmpA_rOmpB; 1.
 KW Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 2249 OUTER MEMBRANE PROTEIN A.
 FT DOMAIN 212 1180 13 X APPROXIMATE TANDEM REPEATS.
 FT REPEAT 212 286 A (TYPE I).
 FT REPEAT 287 358 B (TYPE II).
 FT REPEAT 359 430 C (TYPE II).
 FT REPEAT 431 505 D (TYPE I).
 FT REPEAT 506 577 E (TYPE II).
 FT REPEAT 578 652 F (TYPE I).
 FT REPEAT 653 724 G (TYPE II).
 FT REPEAT 725 799 H (TYPE I).
 FT REPEAT 800 874 I (TYPE I).
 FT REPEAT 875 949 J (TYPE I).
 FT REPEAT 950 1021 K (TYPE II).
 FT REPEAT 1022 1093 L (TYPE II).
 FT REPEAT 1094 1165 M (TYPE II).
 FT REPEAT 1166 1180 TYPE I (INCOMPLETE).
 SQ SEQUENCE 2249 AA; 224333 MW; A9D6646C089DF087 CRC64;
 Query Match 7.0%; Score 121.5; DB 1; Length 2249;
 Best Local Similarity 22.3%; Pred. NO. 8.5;
 Matches 78; Conservative 38; Mismatches 144; Indels 89; Gaps 14;
 QY 14 TNLTNTVVG-----IDSGTIVPHQGYVKLVGFSVPNSAVK-----GDTFKIT 59
 DB 878 TNSLATSISGAGTATLGAVKATTKLTNAASVLTLTNANAVLTGADNTTGGDNVGV- 936
 QY 60 VPKEINLNGV-----TSTAKVPPIAGDQVLAVGIDSDGNVIYTFDYVNTKDD 109
 DB 937 ----LNLGALISQVTDGIGNTSLATISVGAGTATLGAVIKA-----TTTKLTDAASA 986
 QY 110 VKAT--LTPPAYIDPEN-----VKKTGNVTLATIGSTTANKTV----- 146

DB 987 VKETNPVVVVGADNTGNANNGIVTFTGNSTVGNVGTNALTATVNGAGLLQVGGVVK 1046
 QY 147 ----LDYKYEKGFYN--LSIKGTIDDOIKTNN-----TYROTIVNPSGDNIAPVLGT 195
 DB 1047 ANTINLTNASAVTFTNPVVVVGADNTGNANNGIVTFTGNSTVGNVGTNALTATVNGV 1106
 QY 196 NLKPNTDSNALIDQQNTSIKVKVDNAADLSSEYFNP-----ENFEDVTNSVNIPTP 248
 DB -1107 -----AGLLQVGGVGVKANTINLTNANASAVTFTNPVVVGTGADNTGNANNGI-VTFT 1157
 QY 249 NPNQYKVEFNPDDOITTPYIVVNGHIDPNKSGDLALRSTLYGYSNI 297
 DB 1158 GNSTVTGIDGNTNALTAT----VNVGAGITLQAGGSLAANNIDFGARSTL 1202
 RESULT 6
 GP10_DICDI
 ID GP10_DICDI STANDARD; PRT: 544 AA.
 AC Q06885;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Glycoprotein GP100 precursor (P29F8).
 GN GP100.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94124612; PubMed=8294503;
 RA Barth A., Mueller-Taubenberger A., Taranto P., Gerisch G.;
 RA "Replacement of the phospholipid-anchor in the contact site 'A' glycoprotein of D. discoideum by a transmembrane region does not impede cell adhesion but reduces residence time on the cell surface.";
 RT J. Cell Biol. 124:205-215(1994).
 RL -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DEVELOPMENTAL STAGE: RESTRICTED TO THE AGGREGATION STAGE OF DEVELOPMENT IN D.DISCOIDEUM. NO DETECTABLE ACTIVITY IN CELL ADHESION.
 CC -!- PTM: N- AND O-GLYCOSYLATED.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L04286; AAC37369.1;
 DR DictyDb; DD02035; gppA.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 544 GLYCOPROTEIN GP100.
 FT DOMAIN 20 489 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 490 510 POTENTIAL.
 FT DOMAIN 511 544 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 117 208 THR/PRO-RICH.
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 224 224 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 308 308 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 366 366 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 478 478 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 544 AA; 59156 MW; 9B2C2CDE7C7D0AA0 CRC64;
 Query Match 6.9%; Score 119.5; DB 1; Length 544;
 Best Local Similarity 20.3%; Pred. No. 1.8;

Matches 70; Conservative 47; Mismatches 104; Indels 123; Gaps 16;

QY 11 TDITN-QLTNTVIGDSCTTVPHQAGVVKLYGFSVPNSAVKGTFTKTPKELNLNGV 69
 DB 226 TDIEDIEGKTIKYSKQSSSKH-----FQNSIHK-----SI 259

QY 70 TSTAKVPPIMAGDOVLANGVIDSGNVIYFTDYVNTKDDVKATLTPAYIDPENVKKTG 129
 DB 260 ERTNSV-----GGVFEKGSFFYT-----DDYEPSI-----KIG 288

QY 130 NVTLATGIGSTANKTVLVDYKFKYFNLSIKGTIDQITNTYRTIYVNSGDNVI 189
 DB 289 NKICETLSSSTIRCYLTNGTGGC--YTITIDNLLNPIDNNGSNLTICYANPIIDKVI 346

QY 190 A-----PVLGTNL-----KPTND-----SNALIDQNTSIKVKVDNAADLSE 227
 DB 347 GYKDKKLTLLIIGKFNFNATVIEKPNKRNCSNLLSTDLFI-----CELSK 398

QY 228 SYFV--NPENFEDVTSNITFPNPQYKVFENTPDDQITPIYIVVNGHI----- 276
 DB 399 SYHTLSSPK-----TTNTMLSFN-----NSPSTVSEIEMISSYFQIKISTKPS 446

QY 277 ---DPNSKGLDALSTLYGYSNIWRSMSNDNEAFNNGSGG 317
 DB 447 TTDDNNKNDGDSIDSVGKSAVDSSKS-----NNNSGGG 483

RESULT 7
 PM11_CHLPN
 ID PM11_CHLPN STANDARD; PRT; 928 AA.

AC O86164; Q9K299;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable outer membrane protein pmpl1 precursor (Polymorphic membrane protein 11) (Outer membrane protein 4).
 GN PMPL1 OR OMP4 OR CPN0449 OR CP0302.
 OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029/VR-1310;
 RX MEDLINE=20007584; PubMed=10539856;
 RA Christiansen G., Boesen T., Hjerno K., Dagaard L., Mygind P., Madsen A.S., Knudsen K., Falk E., Birkelund S.;
 RT "Identification of two novel genes encoding 97- to 99-kilodalton outer membrane proteins of Chlamydia pneumoniae.";
 RL Infect. Immun. 67:375-383(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VR1310;
 RX MEDLINE=20007584; PubMed=10539856;
 RA Christiansen G., Boesen T., Hjerno K., Dagaard L., Mygind P., Madsen A.S., Knudsen K., Falk E., Birkelund S.;
 RT "Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity.";
 RL Am. Heart J. 138:S491-S495(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RX MEDLINE=9920606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W., Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,

Eisen J., Fraser C.M.;
 "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
 Nucleic Acids Res. 28:1397-1406(2000).
 [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ishii K., Hattori M., Kunara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.";
 Nucleic Acids Res. 28:2311-2314(2000).
 CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES) (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL; AJ001311; CA04672.1;
 DR EMBL; AJ133034; CAB37072.1;
 DR EMBL; AE001628; AAD18593.1;
 DR EMBL; AE002192; AAF38159.1; ALT_INIT.
 DR EMBL; AP002546; BAA98638.1;
 DR PHCI-2DPAGE; O86164;
 DR TIGR; CP0302;
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR InterPro; IPR003357; OMP.
 DR Pfam; PF02385; OMP; 1.
 DR Pfam; PF02415; DUF145; 2.
 KW Outer membrane; Signal; Multigene family; Complete proteome.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 928 PROBABLE OUTER MEMBRANE PROTEIN PMPL1.
 SQ SEQUENCE 928 AA: 98903 MW: 788BCDD62C911402 CRC64;

Query Match 6.7%; Score 115.5; DB 1; Length 928;
 Best Local Similarity 19.3%; Pred. NO. 6.3;
 Matches 73; Conservative 48; Mismatches 137; Indels 121; Gaps 14;

QY 22 VGIDSGTIVPHQAGYKLYGFSVPNSAV-----KGTFTKTPKELNLNGVTSTA 73
 DB 459 LGMDSGTTLSTAGSITITNLGINVDSGLKOPVSLTAKGASNKVIVSGKLNLDIEGNI 518

QY 74 KVPPIAGDQV--LANGVIDSDGNVIYFTDYVNTKDDVKATLTPAYIDPE----- 123
 DB 519 YESHMFSDQLFSLKLTIVDAD-----VDTNVDISSLIPVPAE-DPNSYGFQGO 567

QY 124 -----NVKKTGNVTLATIGTGSTANKTVLVDYKFKYFNLSIKGTIDQIDKTN 172
 DB 568 WNVNWTDTATNTKEATATKTCFVSPERKSALVCNTLWCVFTDIRSLQOLVEIGATG 627

QY 173 NTRYQTIVY-----NPSGDNVIAPVLTGNLKP----- 199
 DB 628 MEHKGQFWSSMTNLFHKTGDNKRKGRHTSGGVIG----GSAHTPKDDLFTFAFCHLF 683

QY 200 NTDNALIDQNTSIKVKVDNAADLSYFVNPENF-----EDVTNSVNITFPNP 250
 DB 684 ARDKDCFIAHNS--RTY--GCTLFFKHSHLTQPNYLRGKRAFESEAIKFFPEIDLA 739

QY 251 NOYKVFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIWRSMSNDNEAF 310
 DB 740 LDQVQFSFSDNRMRMETHYTSL-----PESEG-----SWSNECI- 772

QY 311 NNGSGSGDGIDKPVVPEOP 329
 DB 773 ----AGGIGLDLPFVLSNP 787

or send an email to license@isb-sib.ch.

CC -----

DR EMBL: M37647; AAA26390.1; ALT_INIT.

DR EMBL: AF161079; AAD42234.1; -.

DR EMBL: AJ235273; CAAL15140.1; -.

DR InterPro: IPR003858; rompA_rOmpB.

DR Pfam: PF02708; rompA_rOmpB; 1.

KW Antigen: S-layer; Cell wall; Complete proteome.

FT CHAIN 1 1328 120 KDA SURFACE-EXPOSED PROTEIN.

FT CHAIN 1329 1643 32 KDA BETA PEPTIDE.

FT VARIAT 257 257 V -> A (IN STRAIN BREINL).

FT VARIAT 1010 1010 Y -> D (IN STRAIN BREINL).

FT VARIAT 1450 1450 A -> S (IN STRAIN BREINL).

FT CONFLICT 178 179 AA -> VC (IN REF. 1).

FT CONFLICT 191 201 TTQAPLTGA -> INSKSSYHLVS (IN REF. 1).

FT CONFLICT 212 212 T -> I (IN REF. 1).

FT CONFLICT 313 313 Q -> L (IN REF. 1).

FT CONFLICT 1104 1104 D -> G (IN REF. 2).

FT CONFLICT 1123 1123 T -> S (IN REF. 2).

SQ SEQUENCE 1643 AA; 169854 MW; 735PDF392EG346CC CRC64;

Query Match 6.6%; Score 114.5; DB 1; Length 1643;

Best Local Similarity 21.0%; Pred. No. 15;

Matches 80; Conservative 48; Mismatches 140; Indels 113; Gaps 17;

QY 8 AAGTDITNQLTNVT---VGIDSGTIVYVPHQAGYVKLYGFSVPNSAVKG---DTFKITVP 61

DB 127 AAG-----KILNITCGITVOEASNTINAQNALTKVHGGAANANDLSGLSITF-AAAP 180

QY 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSGNVITYFTD-YVNTKDDVKATL----- 114

DB 181 SVLEFNLIPTQEAPLTLG---ANSKIVNGGCTLNITNGFIQVSDNTFAGIKTINID 236

QY 115 -----TMPAYIDPENVKGTNVTLATGCGSTTANKTVLVVDYKFKYFNLSTKGTID 166

DB 237 DCQGLMFNSTPDAANTLNLQVGGTINPFGIDGT--GKLVLV--SKNGAAEFNVTGTL- 291

QY 167 QIDKTNTNTYRTIYVNPSPGNVIAPLVTLGNLKPNTDSNA--LIDQONTSIKVKYVDNAAD 224

DB 292 -----GGN-----LKGIIELNLAAGKLSOGGAANAVIGTDNCAG 328

QY 225 LSESYFVNPE-----NFEDVTNSV 243

DB 329 RAAGFIVSDVNGNAATISGVYAKNMVIOANAGQVTFEIVDVLGGLGTTNFKTADSKV 388

QY 244 NITPENPQYKVEFTPDQITTPYVIVVNGHI--DPNSKGDALRLSTLYGVNSNIWRS 301

DB 389 IIT-ENSNGSTNFGNLDQIIVPDTKILKGNFIDGVKNNGNTA---GVITFNANGALVS 444

QY 302 MSWDNEVAFNN-----GSGSG 317

DB 445 ASTDPNIATVTNINAIEAGAG 465

RESULT 9

Y013_BPL2 STANDARD; PRT; 738 AA.

AC P42548;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE Hypothetical 81.3 kDa protein precursor (ORE13).

OS Bacteriophage L2.

OC Viruses; dsDNA viruses, no RNA stage; Plasmaviridae; Plasmavirus.

OX NCBI_TaxID=46014;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94215869; PubMed=8163159;

RA Maniloff J., Kambo G.J., Dascher C.C.;

RT "Sequence analysis of a unique temperature phase: mycoplasma virus

RT L2.";

RL Gene 141:1-8(1994).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

OMP_RICPR STANDARD; PRT; 1643 AA.

AC O53020; O3ZCWO; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Outer membrane protein B precursor (168 kDa surface-layer protein)

DE (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)

DE (rOmp B) (Contains: 120 kDa surface-exposed protein (Surface protein

DE antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide).

GN OMPB OR SPAP OR SPA OR RP704.

OS Rickettsia prowazekii.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.

OX NCBI_TaxID=782;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RP STRAIN=Breini;

RP MEDLINE=91045972; PubMed=2122457;

RA Carl M., Dobson M.E., Ching W.M., Dasch G.A.;

RT "Characterization of the gene encoding the protective paracrystalline-

RT surface-layer protein of Rickettsia prowazekii: presence of a

RT truncated identical homolog in Rickettsia typhi.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:8237-8241(1990).

RN [2]

RP SEQUENCE FROM N.A.

RP STRAIN=Breini;

RA Moron C.G., Iu X.J., Walker D.H.;

RT "Sequence analysis of ompB of Rickettsia prowazekii.";

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RP STRAIN=Madrid E;

RX MEDLINE=99039499; PubMed=9823893;

RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,

RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,

RA Eriksson A.-S., Winkler H.H., Kurland C.G.;

RT "The genome sequence of Rickettsia prowazekii and the origin of

RT mitochondria.";

RL Nature 396:133-140(1998).

RN [4]

RP PARTIAL SEQUENCE.

RP STRAIN=Breini;

RX MEDLINE=92114896; PubMed=1370573;

RA Ching W.M., Carl M., Dasch G.A.;

RT "Mapping of monoclonal antibody binding sites on CNBR fragments of

RT the S-layer protein antigens of Rickettsia typhi and Rickettsia

RT prowazekii.";

RL Mol. Immunol. 29:95-105(1992).

RN [5]

RP CLEAVAGE SITE.

RP MEDLINE=92104668; PubMed=1729180;

RA Hackstadt T., Messer R., Clepiak W. Jr., Peacock M.G.;

RT "Evidence for proteolytic cleavage of the 120-kilodalton outer

RT membrane protein of rickettsiae: identification of an avirulent mutant

RT deficient in processing.";

RL Infect. Immun. 60:159-165(1992).

CC -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR

CC STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL

CC VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.

CC -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR.

CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-

CC LAYER WITH HEXAGONAL SYMMETRY.

CC -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: L13696; AAA87969.1; -
 KW Hypothetical protein; Signal; Transmembrane.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 738 HYPOTHETICAL 81.3 KDA PROTEIN.
 FT TRANSMEM 612 632 POTENTIAL.
 FT TRANSMEM 712 732 POTENTIAL.
 SQ SEQUENCE 738 AA; 81312 MW; 87479A3267C89846 CRC64;

Query Match 6.6%; Score 114; DB 1; Length 738;

Best Local Similarity 22.4%; Pred. NO. 5.8;
 Matches 75; Conservative 50; Mismatches 110; Indels 100; Gaps 19;

41 NYGESVPSNAVKDFKLTVPKELNLTSTAKVPPIMAGDOV---LANGVIDS---DGN 95
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 420 NDGYSPTVGVNVEFSVT-----NSNGQTSSI-IAPVHVVDIVNPVINGSDTVHISYD 473
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 96 VIYTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGTSTANKTVLVDYKEYGK 155
 :
 474 QTENVTVNVS-----LVSDNY-----TGLSISIKENTYVKNKLG 513
 :
 156 FYNLSIKTIDIDKTN--TYRQTIYVNP-----SGDNVIAPVLTGN----- 196
 :
 514 Y-----KITVOADPSGNGITLRTIVVNDGIGPVFNGINTITASINENITVEQIKAGLA 568
 :
 197 ---LKPNTDSNALDOON-----TSIKVYK-----VDNAAD-----LSSEYFVNPENFE 237
 :
 569 AIDAIDGNVTVSIVVDSNLGKANTGVYEVFRVDAAGNQTFHTVTVSIVASPPGFY 628
 :
 238 DY-TNSVNITFPNP-----QYKVFENTPD-DQITTPYIVVYNGHIDPNKGLDALRSLTY 291
 :
 629 ILNSNSVRL-LPGANLTIEQILNINLASDAENISTNYTVSPGIYN-----LSFTLY 679
 :
 232 GYNSNIWRSWHDNEVAFNNGSGGIDKPVVP 326
 :
 680 GESHOVSITVLGONDSI-----IPTPVP 703

RESULT 10

SLAP-BACST STANDARD; PRT: 1228 AA.
 P35825;
 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE S-layer protein precursor (Surface layer protein).
 GN SBSA.
 OS Bacillus stearothermophilus.
 OC Bacteria; Firmicutes; Bacillales; Geobacillus.
 OX NCBI_TaxID=1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PV72;
 RX MEDLINE=943320770; PubMed=8045409;
 RA Kuen B., Sleytr U.B., Lubitz W.;
 RT "Sequence analysis of the sbsA gene encoding the 130-kDa
 RT surface-layer protein of Bacillus stearothermophilus strain PV72.";
 RL Gene 145:115-120(1994).
 CC -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
 CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
 CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
 CC S-LAYER WITH HEXAGONAL SYMMETRY.
 CC -----

* This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: X71092; CAA50409.1; -
 DR EMBL: AX000218; CAB77069.1; -
 KW PIR: S34365; S34365.
 KW Signal; Cell wall; S-layer.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 1228 S-LAYER PROTEIN.
 SQ SEQUENCE 1228 AA; 131076 MW; 9F1383AD810C0B0F CRC64;

Query Match 6.6%; Score 113.5; DB 1; Length 1228;

Best Local Similarity 22.1%; Pred. NO. 12;
 Matches 90; Conservative 51; Mismatches 139; Indels 127; Gaps 24;

QY 12 DITNQLTNTVTVG-----IDSGTTVVPH-----QAGYVKLNIGFSV---PNS 49
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 769 DAVTTLTNVDAGQKFTIOFSEBLKTSGLVGGKVTVEKLTNNGWVDAGTGTTVSVAPKT 828
 :
 50 AVKGDFTKITVPKELNLTSTAKVPI-MAGDOVLANGVIDSDGNYI-----YTPTD 102
 :
 829 DANG---KVTA-AVVTLTGLDNNDKAKLRLVVDKSDTDGIADVAGNVIKEKDILIRYS 884
 :
 103 YVNT-----KDDVKATLTMP-----AYIDPENV---KKTGN 130
 :
 885 WRHTVASVKAADKDGQNASAAFTPTSTAIDTTKSLLEVNETDLAEVKPENIVVKNDAAGN 944
 :
 131 VLTATGI---GSTTANKTVLVDYE--KYGKFYNLSIKGTIDQIDKTNNTY---ROTIVN 182
 :
 945 AVAGTVTALDGST--NKFVFTPSQELKAGTVSVTIDGVRDKVGTISKYITSKTVTSAN 1002
 :
 183 PSGDNVIAPVLTGNLKNPTDSNALIDQO-----NTSIKYKVDNAADLSSEYFVNPENF 236
 :
 1003 PT-----LSSSIADGAVNVDRSKTITIEFSDVSNPTITLKKADG----- 1043
 :
 237 EDVTVSNVITFPNQ-----YKVFEN---TPDQITTPYIVVNGHI-----DPNSK----- 281
 :
 1044 ---TSFTNTVLNVNNEKNKYIVFHKGVTIDE--FTQYELAVSKDFOTGTDIDSKVTFI 1098
 :
 282 -GDLA---LRSTLYGYNISIIWRSNMDNEVAFNNGSGGIDKPV 324
 :
 1099 TGSVATDEVKPAKLVCGS---WNGTSYTDAAATRLRSVADFVAEPV 1142

RESULT 11

YFAL-ECOLI STANDARD; PRT: 1250 AA.
 AC P45508; P45507; P45506; P39441; P76468; P77487;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yfaL precursor.
 GN YFAL OR B2233.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;

RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sappei G., Satoh Y., Sivasundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takeuchi K., Uehara K., Wada C.,
RA Yamagata S., Horuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN [3]
RP PRELIMINARY SEQUENCE OF 1-938 FROM N.A.
RX MEDLINE=84272624; PubMed=6087316;
RA Carlson J., Fuchs J.A., Messing J.;
RT "Primary structure of the Escherichia coli ribonucleoside diphosphate
RT reductase operon.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4294-4297(1984).
RN [4]
RP SEQUENCE OF 925-1198 FROM N.A.
RX STRAIN=K12 / EMG2;
RA Estep P., O'Keefe T., Robinson K., Church G.M.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1180-1250 FROM N.A.
RX STRAIN=OV6;
RA MEDLINE=88201664; PubMed=2834621;
RA Hussain K., Elliott E.J., Salmund G.P.C.;
RT "The parD-mutant of Escherichia coli also carries a gyrAam mutation.
RT The complete sequence of gyrA.";
RL Mol. Microbiol. 1:259-273(1987).
RN [6]
RP IDENTIFICATION.
RX MEDLINE=96032851; PubMed=7567469;
RA Borodovsky M., McIninch J., Koonin E.V., Rudd K.E., Medigue C.,
RA Danchin A.;
RT "Detection of new genes in a bacterial genome using Markov models for
RT three gene classes.";
RL Nucleic Acids Res. 23:3554-3562(1995).
CC -!- SIMILARITY: TO E.COLI YDEK.
CC -!- CAUTION: REF.3 SEQUENCE DIFFERS EXTENSIVELY FROM THAT SHOWN BY
CC MANY FRAMESHIFTS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000313; AAC75293.1; -;
DR EMBL; D90855; BAA16052.1; ALT_INIT.
DR EMBL; D90854; BAA16050.1; ALT_INIT.
DR EMBL; K02672; -; NOT_ANNOTATED_CDS.
DR EMBL; U30459; AAA74094.1; -;
DR EMBL; Y00544; -; NOT_ANNOTATED_CDS.
DR EcoGene; EG12850; yfaL.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; DUF145; 1.
DR Pfam; PF03212; Pertactin; 1.
DR Hypothetical protein; Repeat; 1.
KW SIGNAL 1 23
FT CHAIN 24 1250
FT DOMAIN 919 948
FT CONFLICT 28 30
FT CONFLICT 40 40
FT CONFLICT 65 66
FT CONFLICT 431 431
FT CONFLICT 433 434
FT CONFLICT 478 478
FT CONFLICT 773 773

FT CONFLICT 853 853 V -> M (IN REF. 3).
FT CONFLICT 923 924 PP -> AT (IN REF. 3).
FT CONFLICT 948 994 PAYQVULNAKVGGLNLRANQAQFMRDRDHAGGQGTIN
FT SEQUENCE 1250 AA; 131152 MW; 17F98C05E299FC95 CRC64;
SQ
Query Match 6.6%; Score 113.5; DB 1; Length 1250;
Best Local Similarity 21.9%; Pred. No. 12;
Matches 90; Conservative 45; Mismatches 137; Indels 139; Gaps 19;
QY 4 ADAPAAAGTDTITNQLTN---VTVGIDS-----GTTVYPHQAGYVKL----- 40
Db 618 SDAAGHGRIEMRADGEVAVDAGVDTQWQALMADSSGQHQDEGSTLTKTGAGTLELTASG 677
QY 41 --NYGFSVPNSAVKGDFTKFTVPKELNL---NGVTSTAKVPPIMAGDQVLANGVID-SDG 94
Db 678 TTQSAVRVEEGTLKGDVADI-LPYASSLWGGGATFVTGADQDIQSIDAISSTGTDISDG 736
QY 95 NVI-YTFETD-----YVNTKDDYKATLTMPAYIDPE-----NVKKTGNVT 132
Db 737 TVLRLTGQDTSVALNASLFPNGDGTLLVNAVDTGVTLTGELNTNLETDSLTYSLVNVVNGNLT 796
QY 133 LATGI-----GTTANKTVLVDYKGYKFNLSIKGTIDQIDKTNNTYRQ 177
Db 797 NTSGAVSLQGVAGDTLTVNGDYGTTGGTLLDSELNGD-----DSVSDQLVWNGNTAGN 850
QY 178 TIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKYVDNAD-----LS 226
Db 851 T-----TVVNSITGIGEPT-----STGIKV--VDFAADPTQFONNAQFSLA 890
QY 227 ESYFVN-----PENFEDVTSNVT-----PPNPQYKVEENTPDQITTPY 268
Db 891 GSGYVNMGAYDYLVEDNNDWYLRSGEVTPPSPDPDPDPDPDPDPDPDPDPDPPTPAY 950
QY 269 IYVNVGHIDPNKGDALRLSTLYGNSNIWMSWDNVAFNNGSGSDG 319
Db 951 QPVLNKVG-----GYLNNLRANQAQFMME---RRDHAGG 984

RESULT 12
PIP_LACLC STANDARD; PRT; 1902 AA.
AC PIP_LACLC
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE PI-type proteinase precursor (EC 3.4.21.-) (Wall-associated serine
DE proteinase).
GN PRTP.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OG Plasmid pW05.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1359;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=WG2;
RX MEDLINE=88149035; PubMed=3278687;
RA Kok J., Leenhouts K.J., Haandrikman A.J., Ledeboer A.M., Venema G.;
RT "Nucleotide sequence of the cell wall proteinase gene of
RT Streptococcus cremoris Wg2.";
RL Appl. Environ. Microbiol. 54:231-238(1988).
CC -!- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE
CC GROWTH OF THE BACTERIA ON MILK.
CC -!- CATALYTIC ACTIVITY: Endopeptidase activity with very broad
CC specificity, although some substrate preference have been noted,
CC e.g. large hydrophobic residues in the P1 and P4 positions, and
CC pro in the P2 position. Best known for its action on caseins,
CC although it has been shown to hydrolyse hemoglobin and oxidized
CC insulin B-chain.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (potential).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M24767; AA17677.1; -
CC HSP: P00782; LS01.
CC MEROPS: S08.019; -
CC InterPro: IPR001899; Gram_pos_anchor.
CC InterPro: IPR003137; PA.
CC InterPro: IPR000209; Peptidase_S8.
CC Pfam: PF00082; Peptidase_S8; 3.
CC Pfam: PF00746; Gram_pos_anchor; 1.
CC Pfam: PF02225; PA; 1.
CC PRINTS: PR00723; SUBTILISIN.
CC TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
CC PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
CC PROSITE: PS00136; SUBTILASE_ASP; 1.
CC PROSITE: PS00137; SUBTILASE_HIS; 1.
CC PROSITE: PS00138; SUBTILASE_SER; 1.
CC Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
KW Signal; Plasmid.
FT SIGNAL 1 33
FT PROPEP 34 187
FT CHAIN 188 1870
FT PROPEP 1871 1902
FT ACT_SITE 217 217
FT ACT_SITE 281 281
FT ACT_SITE 620 620
FT SITE 1867 1871
FT MOD_RES 1870 1870
FT SEQUENCE 1902 AA; 199910 MW; 2901C7F19B2E5D0B CRC64;

Query Match 6.6%; Score 113.5; DB 1; Length 1902;
Best Local Similarity 21.1%; Pred. No. 21;
Matches 71; Conservative 35; Mismatches 100; Indels 131; Gaps 14;

QY 3 AADAPAGDTINQ-----LTNVTGIDSGT-----TYPHQGVVVL- 40
DB 506 AQAAGAAGLIIVNDGTAPFVSMALTTFTFFGLSSVTGQKLVDMVTAHPDDSLGVKTA 565
QY 41 -----NYGFSVPNSAVKGTFTKIVP-----RELNLNGVTS---TAK 74
DB 566 LTLVPNQKYTEDKMSDFTSYG-PVSNLSFKPD---ITAPGGNIWSTQNNNGYTNKSGTSM 621
75 VPPIMAGDQVLANGVIDSGNVIYTF-----TDYVNTKDDVKATLTPAYIDPENY 125
622 ASPFIAGSOALLKQALNNKNNPFYAYIKLKGATLTDFLKT-----VEMNTA 668
QY 126 KKTGNVTLATGISTANKTVLDVYKFKYNSLTKGTIDQDK----- 170
DB 669 QPINDINYNVIVSPROGAGLV-----VKAIDAILEKNPSTVWAENGYPAVE 717
QY 171 -----TNNYRTQTYVPSGDNVIAPIVLTGLNKPTDSNAL-----IDQNTSIVKYVD 220
DB 718 LKDFSTSTDKTKLTF-----INSITHETLYQWDSNTDPAVYTSATDPSNGVLVDKKID 771
QY 221 NADLSESYFVNPENFEDVTSNVTFFPNPNQYKVEF 257
DB 772 GAA-----IKAGSNITVPAGKTAQIEF 793

RESULT 13

P2P_LACPA
ID P2P_LACPA STANDARD; PRT; 1902 AA.
AC Q02470;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE PII-type proteinase precursor (EC 3.4.21.96) (Lactocepin) (Cell wall-

DE associated serine proteinase) (LP151).
GN PRT.
OS Lactobacillus paracasei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1597;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCDO 151;
RX MEDLINE=92381481; PubMed=1512565;
RA Holck A., Naes H.;
RT "Cloning, sequencing and expression of the gene encoding the cell-
RT envelope-associated proteinase from Lactobacillus paracasei subsp.
RT paracasei NCDO 151.";
RL J. Gen. Microbiol. 138:1353-1364(1992).
RN [2]
RP SEQUENCE OF 189-196.
RX MEDLINE=92226694; PubMed=1564442;
RA Naes H., Nissen-Meyer J.;
RT "Purification and N-terminal amino acid sequence determination of the
RT cell-wall-bound proteinase from Lactobacillus paracasei subsp.
RT paracasei.";
RL J. Gen. Microbiol. 138:313-318(1992).
CC -!- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE
CC GROWTH OF THE BACTERIA ON MILK.
CC -!- CATALYTIC ACTIVITY: Endopeptidase activity with very broad
CC specificity, although some subsite preference have been noted,
CC e.g. large hydrophobic residues in the P1 and P4 positions, and
CC Pro in the P2 position. Best known for its action on caseins,
CC although it has been shown to hydrolyse hemoglobin and oxidized
CC insulin B-chain.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M83946; AAA25248.1; -
CC PIR: B44858; B44858.
CC HSP: P00782; LS01.
CC MEROPS: S08.019; -
CC InterPro: IPR001899; Gram_pos_anchor.
CC InterPro: IPR003137; PA.
CC InterPro: IPR000209; Peptidase_S8.
CC Pfam: PF00082; Peptidase_S8; 3.
CC Pfam: PF00746; Gram_pos_anchor; 1.
CC Pfam: PF02225; PA; 1.
CC PRINTS: PR00723; SUBTILISIN.
CC TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
CC PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
CC PROSITE: PS00136; SUBTILASE_ASP; 1.
CC PROSITE: PS00137; SUBTILASE_HIS; 1.
CC PROSITE: PS00138; SUBTILASE_SER; 1.
CC Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
KW Signal.
FT SIGNAL 1 33
FT PROPEP 34 187
FT CHAIN 188 1870
FT PROPEP 1871 1902
FT ACT_SITE 217 217
FT ACT_SITE 281 281
FT ACT_SITE 620 620
FT SITE 1867 1871
FT MOD_RES 1870 1870
FT SEQUENCE 1902 AA; 200253 MW; D8C9F38CEE5DA582 CRC64;

Query Match

6.6%; Score 113.5; DB 1; Length 1902;

Best Local Similarity 23.28; Pred. No. 21;		Matches 76; Conservative 35; Mismatches 104; Indels 113; Gaps 17;	
Qy	3	AADAPAAAGTDTITNQ-----LTNV-----TVGIDSGT-----TVYPHOAGYVKL- 40	
-Db	506	AQAAGAGLIIVNDGTATPLTSLRLLTPTTGLSSKTQKLVWDVTAHPDDSLGKVIA 565	
Qy	41	-----NYGFSVPNSAVKGDFTKTIIVP-----KELNLGVTS---TAK 74	
-Db	566	LTLLPNKYTEDKMSDFTSYG-PVSNLSFKPD---ITAPGNIWSTONNGYTNMSGTSM 621	
Qy	75	VPIIMAGDOVLANGVIDSDGNVIYFTFTDYVNTKDDVKATLTMPAYIDPENWKTGNVTLA 134	
-Db	622	ASPEIAGSOALLKQALNNKNPFY--ADYKQLKG---TALT-----DFLKTVENTA 668	
Qy	135	TGISTTANKTVLDVYKFKYNSLTKGTIDQIDK-----TNN 173	
-Db	669	QPINDIYN-NVIVSPRQAGL-VDVKAIDALEKNPSTVVAENGYPAVELKDTSTDK 726	
Qy	174	TYRQTIYVNPSSGDNVIAPVLTKNLKPNNTDSNAL-----IDQNTSIKVKYVDNAADLSESY 229	
-Db	727	TFKLTF-----TNRTHLTYQMSNTDTNAVYTSATDPNSGVLYDKKIDGAA----- 774	
Qy	230	FVNPENFEDVTSNVTFFPNPNQYKVEF 257	
-Db	775	-----IKAGSDITVPAGKTAQIEF 793	
RESULT 14			
P3P_LACLC	STANDARD; PRT: 1902 AA.		
ID	P3P_LACLC		
AC	P15292;		
DT	01-APR-1990 (Rel. 14, Created)		
DT	01-APR-1990 (Rel. 14, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Pili-type proteinase precursor (EC 3.4.21.96) (Lactococpin) (Cell wall-associated serine proteinase).		
GN	PTTP.		
OS	Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).		
OG	Plasmid.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.		
OX	NCBI_TaxID=1359;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 188-197.		
RC	STRAIN=SK11;		
RX	MEDLINE=89340435; PubMed=2760036;		
RA	Vos P., Simons G., Siezen R.J., de Vos W.M.;		
RE	"Primary structure and organization of the gene for a procaryotic, cell envelope-located serine proteinase.";		
RE	J. Biol. Chem. 264:13579-13585(1989).		
CC	-1- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE GROWTH OF THE BACTERIA ON MILK.		
CC	-1- CATALYTIC ACTIVITY: Endopeptidase activity with very broad specificity, although some substrate preference have been noted, e.g. large hydrophobic residues in the P1 and P4 positions, and pro in the P2 position. Best known for its action on caseins, although it has been shown to hydrolyse hemoglobin and oxidized insulin B-chain.		
CC	-1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (Potential).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; J04962; AAA03533.1; ALT_SEQ.		
DR	PIR; A32634; A32634.		
DR	HSSP; P00782; 2S8T.		

MEMBERS; S08.019; -			
DR InterPro; IPR001899; Gram_pos_anchor.			
DR InterPro; IPR001137; PA.			
DR InterPro; IPR000209; Peptidase_S8.			
DR Pfam; PF00082; Peptidase_S8; 3.			
DR Pfam; PF00746; Gram_pos_anchor; 1.			
DR Pfam; PF02225; PA; 1.			
DR PRINTS; PR00723; SUBTILISIN.			
DR TIGRFS; TIGR01167; LPXTG_anchor; 1.			
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.			
DR PROSITE; PS00136; SUBTILASE_ASP; 1.			
DR PROSITE; PS00137; SUBTILASE_HIS; 1.			
DR PROSITE; PS00138; SUBTILASE_SER; 1.			
KW Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;			
KW Signal; Plasmid.			
FT SIGNAL	1 33		
FT PROPEP	34 187		
FT CHAIN	188 1870		
FT PROPEP	1871 1902		
FT ACT_SITE	217 217		
FT ACT_SITE	281 281		
FT ACT_SITE	620 620		
FT SITE	1867 1871		
FT MOD_RES	1870 1870		
SQ SEQUENCE	1902 AA; 200550 MW; 87CECBAA9345F9D3 CRC64;		
Query Match 6.6%; Score 113.5; DB 1; Length 1902;			
Best Local Similarity 21.08; Pred. No. 21;			
Matches 70; Conservative 38; Mismatches 101; Indels 125; Gaps 15;			
Qy	3	AADAPAAAGTDTITN-----QLTNV-----TVGIDSGT-----TVYPHOAGYVKLN 41	
-Db	506	AQAAGAGLIIVNDGTATPMTSIALTTPTTGLSSVTGKLVWDVTAHPDDSLGVKIT 565	
Qy	42	YGSFVSPNSAVKGD-----TFK---ITVP-----KELNLGVTS---TAKVPP 77	
-Db	566	LAM-LPNQKYTEDKMSDFTSYGVPVSNLSFKPDITAPGGNIWSTONNGYTNMSGTSMA 624	
Qy	78	IMAGDOVLANGVIDSDGNVIYTF-----TDYVNTKDDVKATLTMPAYIDPENWKT 128	
-Db	625	FIAGSOALLKQALNNKNPFYAYKQLKGTALTDFLKT-----VEMTAQPI 671	
Qy	129	GNVTLATGISTTANKTVLDVYKFKYNSLTKGTIDQIDK----- 170	
-Db	672	NDINYNVIVSPRQAGLVD-----VKAIDALEKNPSTVVAENGYPAVELKD 720	
Qy	171	---TNNTYRQTIYVNPSSGDNVIAPVLTKNLKPNNTDSNAL-----IDQNTSIKVKYVDNA 223	
-Db	721	FTSTDKTFKLTF-----TNRTHLTYQMSNTDTNAVYTSATDPNSGVLYDKKIDGAA 774	
Qy	224	DLSEYFVNPENFEDVTSNVTFFPNPNQYKVEF 257	
-Db	775	-----IKAGSNITVPAGKTAQIEF 793	
RESULT 15			
AIDA_ECOLI			
ID	AIDA_ECOLI	STANDARD; PRT: 1286 AA.	
AC	Q03153;		
DT	01-JUN-1994 (Rel. 29, Created)		
DT	01-JUN-1994 (Rel. 29, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Adhesin aidA-I precursor.		
GN	AIDA-I.		
OS	Escherichia coli.		
OG	Plasmid pIB6.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Escherichia.		
OX	NCBI_TaxID=562;		
OX	[1]		
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 50-56.		
RP	STRAIN=O126:H27 / 2787;		
RC	MEDLINE=92326638; PubMed=1625582;		
RX			

Query Match 6.5%; Score 112; DB 1; Length 1902;
 Best Local Similarity 21.1%; Pred. No. 26;
 Matches 60; Conservative 37; Mismatches 90; Indels 98; Gaps 13;

QY 21 TVGIDSGTGVPHQAGYVK-----LNYGFSVPNSAVKGDTEKITVP-----KELNLGV 69
 DB 559 SLGVKIALTLVPNQK-YTEDKMSDFTSYG-PVSNLSFRPD---ITAPGGINWSTONNGY 613
 QY 70 TS---TAKVPPIMAGDQVLANGVIDSDGNVIYTF-----TDVYNTKDDVKATLTMP 117
 DB 614 TNMSGTSMAFPFIAGSQALLKQALNNKPNPFYAYYKQLKGALTDLFKT-----662
 QY 118 AVIDPENVKKTGNVTLATIGIGSTTANKTVLVYDYKYGYFNYSIKGTIDQIDK-----170
 DB 663 --VEMTAQPIINDINYNVIVSPRRQAGLVD-----VRAAIDALEKPNSTVVA 709
 QY 171 -----TNTYROTIIYVNPSPGDVNIAPVLTCNLKPNKPTDSNAL-----IDQNT 212
 DB 710 ENGYPVELKDFTSKTKLTF-----TNTTHLTYQMSDNTDTNAVYTSATDPNSG 763
 QY 213 SIKVYKVDNAADLSSEYFVNPENFEDVTSVNIPTNPNOYKVEF 257
 DB 764 VLYDKKIDGAA-----IKAGSNITVPAGKTAQIEF 793

RESULT 18
 Y109_YEAST
 ID Y109_YEAST STANDARD; PRT; 995 AA.
 AC P40442;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 99.7 kDa protein in SDL1 5' region precursor.
 GN YIL169C OR YI9402.07C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Davellin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 246921; CAA87023.1;
 DR SGD; S0001431; YIL169C.
 DR InterPro; IPR004089; Chmtaxis_transd.
 DR InterPro; IPR000727; T-SNARE.
 DR PROSITE; PS50192; T-SNARE; UNKNOWN_1.
 DR Hypothetical protein; Signal.
 KW SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 995
 FT DOMAIN 92 154
 FT CARBOHYD 28 28
 FT CARBOHYD 35 35
 FT CARBOHYD 468 468
 FT CARBOHYD 664 664
 SQ SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;

Query Match 6.4%; Score 111; DB 1; Length 995;

RESULT 17
 P2P_LACLC
 ID P2P_LACLC STANDARD; PRT; 1902 AA.
 AC P15293;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE PII-type proteinase precursor (EC 3.4.21.96) (Lactocepin) (Cell wall-
 DE associated serine proteinase) (LP151).
 GN PRT.
 OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
 OG Plasmid pLP763.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCDO 763;
 MEDLINE=89313288; PubMed=2501630;
 KIWAKI M., IKEMURA H., SHIMIZU-KADOTA M., HIRASHIMA A.;
 RT "Molecular characterization of a cell wall-associated proteinase gene
 from Streptococcus lactis NCDO763.";
 RL Mol. Microbiol. 3:359-369(1989).
 CC -!- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE
 CC GROWTH OF THE BACTERIA ON MILK.
 CC -!- CATALYTIC ACTIVITY: Endopeptidase activity with very broad
 CC specificity, although some subsite preference have been noted,
 CC e.g. large hydrophobic residues in the P1 and P4 positions, and
 CC pro in the P2 position. Best known for its action on caseins, and
 CC although it has been shown to hydrolyse hemoglobin and oxidized
 CC insulin B-chain.
 CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (Potential).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X14130; CAA32350.1;
 DR PIR; S06997; S06997.
 DR HSSP; P00782; 2S8T.
 DR MEROPS; S08.019;
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR003137; PA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 3.
 DR Pfam; PF00746; Gram_pos_anchor; 1.
 DR Pfam; PF02225; PA; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR TIGRfams; TIGR01167; LPXtg_anchor; 1.
 DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
 KW Signal; Plasmid.
 FT SIGNAL 1 33 POTENTIAL.
 FT PROPEP 34 187
 FT CHAIN 188 1870
 FT PROPEP 1871 1902
 FT ACT_SITE 217 217
 FT ACT_SITE 281 281
 FT ACT_SITE 620 620
 FT ACT_SITE 1867 1871
 FT MOD_RES 1870 1870
 SQ SEQUENCE 1902 AA; 200139 MW; 4B8DB844D88CDF7 CRC64;

Query Match	6.4%	Score 110.5;	DB 1;	Length 1176;	
Best Local Similarity	22.5%;	Pred. NO. 18;			
Matches	72;	Conservative 45;	Mismatches 110;	Indels 93;	Gaps 16

QY	28	TTVYPHOAGYKVLN	YGFSPNSAVKGD	FEKITVPKELNLCV	TSKAVPPIMAGD	--OVL 85
DB	157	STVKPWAKSYLEI	-----AVANGV	KGS-----E	ANGKTNLNP	NAPITRODFAVVF 202
QY	86	ANGVIDSDG	-----NVIYTF	DYVNTKDDVKAT	LTPAYIDPE	-----N 124
DB	203	SRTIENVADATPK	VDKIEVVDAKT	LNVLSDGCTETV	LKALEPNKET	EVTEFKIKDVEYK 262
QY	125	VKKTGNVTLATG	IGSTTAN--KTV	LVDEYKGYKFP	YINLSIKGTIDQ	---IDKTNNTYRQTI 179
DB	263	AKVTYVVTATAT	KAIVSASATNL	KEVVVEED	-----GTVDK	ETAEDAAN----- 304
QY	180	YVNPSGDNVIA	PVLGNLKPNT	SDNALIDQONT	SIKVKYVDNA	ADLSYFVAPENFEDV 239
DB	305	YALKSGKTIK	SVSLAADNK--	TATVTLTDKLN	-----NNKADA	-----ISISNVKAG 349
QY	240	TSNVNITFPNP	QYKVFENTP	DDQITPYIVV	VANGHIDPN	SKGLALRLSTLYGYSNIIW 299
DB	350	DKEINVK-----	NVEFTA	VDNKI--DEV	TEV-----KSLGT	KAVKVTLSPEVNLS 394
QY	300	RMSWDNEVAF	NPN---GSGS	316		
DB	395	TNFTLDGKAY	FCGVVYGAGN	414		

RESULT 20	
OMPB_RICTY	STANDARD; PRT; 1645 AA.
ID	OMPB_RICTY
AC	P96989;
DT	15-JUL-1998 (Rel. 36, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Outer membrane protein B precursor (168 kDa surface-layer protein)
DE	(Surface protein antigen) (Cell surface antigen 5) (Scas) (fOmpB)
DE	(rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE	antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN	OMPB OR SLP.
OS	Rickettsia typhi.
OC	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC	Rickettsiaceae; Rickettsiae; Rickettsia.
NCBI_Taxid=785;	
SEQUENCE FROM N.A.	
STRAIN=	Wilmington;
MEDLINE=	94040787; PubMed=8224886;
Hahn M.-J., Kim K.-K., Kim I., Chang W.-H.;	
"Cloning and sequence analysis of the gene encoding the crystalline	
surface layer protein of Rickettsia typhi.";	
Gene 133:129-133(1993).	
PARTIAL SEQUENCE.	
STRAIN=	Wilmington;
MEDLINE=	92114896; PubMed=1370573;
Ching W.M., Carl M., Dasch G.A.;	
"Mapping of monoclonal antibody binding sites on CNBr fragments of	
the S-layer protein antigens of Rickettsia typhi and Rickettsia	
prowazekii";	
Mol. Immunol. 29:95-105(1992).	
IDENTIFICATION OF CLEAVAGE SITE.	
MEDLINE=	92104668; PubMed=1729180;
Hackstadt T., Messer R., Cleplak W., Peacock M.G.;	
"Evidence for proteolytic cleavage of the 120-kilodalton outer	
membrane protein of rickettsiae: identification of an avirulent	
mutant deficient in processing.";	
Infect. Immun. 60:159-165(1992).	
-1- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR	
STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL	

CC Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;

CC Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;

CC Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;

```

RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitagawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubram S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
CC -!- SIMILARITY: CONTAINS 13 BIG-1 DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE INTIMIN/INVASIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: AE000289; AAC75042.1; ALT_INIT.
DR EMBL: D90837; BAA15800.1; -.
DR EMBL: D90836; BAA15799.1; ALT_INIT.
DR EcoGene: EG13378; yeeJ.
DR InterPro: IPR003344; Big_1.
DR InterPro: IPR003535; Intimin.
DR InterPro: IPR002482; LysM.
DR InterPro: IPR000601; PKD_domain.
DR Pfam: PF02369; Big_1; 13.
DR PRINTS: PR01369; INTIMIN.
DR SMART: SM00257; LysM; 1.
DR SMART: SM00089; PKD; 4.
KW Hypothetical protein; Repeat; Complete proteome.
FT DOMAIN 738 834
FT DOMAIN 840 931
FT DOMAIN 932 1033
FT DOMAIN 1042 1137
FT DOMAIN 1146 1237
FT DOMAIN 1246 1350
FT DOMAIN 1351 1448
FT DOMAIN 1449 1553
FT DOMAIN 1554 1655
FT DOMAIN 1661 1754
FT DOMAIN 1763 1853
FT DOMAIN 1855 1950
FT DOMAIN 1952 2053
FT DOMAIN 2055 2150
FT CONFLICT 105 105 S -> G (IN REF. 2).
SQ SEQUENCE 2358 AA; 248599 MW; 232249750BF631ED CRC64;

Query Match 6.2%; Score 108; DB 1; Length 2358;
Best Local Similarity 19.6%; Pred. No. 62;
Matches 93; Conservative 46; Mismatches 166; Indels 170; Gaps 18;

QY 1 MYADAPAGTDTITNQLTNVTVGIDSGTIVPHQAGYKLVNFGFVSPNSAVKGDFTKITV 60
Db 1879 LTAULTSANGTPEGVQVNFVTPGATL-----SGKVRTNSSGOAPVVLTSNKVCTYTV 1934
QY 61 PRELNNGVT-----STAKVPPTMAGDOVLA-----NGVIDSDGNVI 97

```

```

Db 1935 TASFH-NGVTIQTTQTVKVTGNSSTAHSVAFIADPSTIAATNTDLSTLKAATVEDGSGNLI 1993
QY 98 YFTFDYVNTKDDVKATLTPAVIDPEN-----VK--KTGNVTLA---TGIGSTTANKTV 146
Db 1994 EGLTVYFALKSG-SATLTSLATVTDQNGIATTSVKGAMTGSVTSVAVTTAGMGQTVDTL 2052
QY 147 LV---DYEKYKGFYNLSIKG-----YIVVV-----NGHIDPNS 280
Db 2053 VAGPADTSQSVLKSNRSLKGDYTDSAELRLVLDHISGNPIKVSQEGMEFVQSGTNVPYIK 2112
QY 164 --TIDQIDKNTNTYROTIIYVNPNGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIKVKVD- 220
Db 2113 ISAIYSLNNGDYKATVTGGGEGIATLIPVLNGVHQAGLSTTIQFTRAEKIMSGTQSV 2172
QY 221 NAADLSES-----YFVNPENFE-----DVTNSVNIPTPNPNO 252
Db 2173 NCTDLPTTTTTPSQGTGAYQLNNDNFAPGKTAADYEFSSASWVDVDTGKVTGN--- 2229
QY 253 YKVEFTPDQIT-Tp-----YIVVV-----NGHIDPNS 280
Db 2230 ----VGSNSERITATPKSGGPSYVYVIRVKSWMVNAAGEAFMIYSLAENFCSSNGYV 2285
QY 281 K-----GDALRSTLYGYNIIWRSMNSWDNEVAFNNGSGSD 318
Db 2286 NYLNHCSSRGIGSLYSEMGDMGHYTTDAGFQSNMYKSSSPANSSEQYVYVSLATGD 2340

Search completed: June 23, 2003, 14:02:08
Job time : 18.3245 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:10 ; Search time 42.8422 seconds

(without alignments)
1591.926 Million cell updates/sec

Title: US-10-056-052A-4

Perfect score: 1732

Sequence: 1 MVAADAPAAAGTDITNLTNV.....NGSGGIDGIDKPVVPEQDPDE 331

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

1 number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1727	99.7	933	2	Q53653
2	1517	87.6	935	16	Q932C5
3	1517	87.6	989	16	Q932C5
4	1112	64.2	881	2	Q93MH7
5	416.5	24.0	961	16	Q93RD3
6	409	23.6	940	2	Q53682
7	353	20.4	931	2	Q9K113
8	349	20.2	1092	2	Q70022
9	324.5	18.7	1038	16	Q99RD2
10	323.5	18.7	1166	2	Q86489
11	320.5	18.5	1141	16	Q99W46
12	320.5	18.5	1141	16	Q932F7
13	312.5	18.0	877	16	Q99R07
14	312.5	18.0	913	2	Q86476
15	298	17.2	1171	2	Q9KWX6
16	257.5	14.9	1315	2	Q86488

17	240.5	13.9	1385	16	Q99W47	Q99W47 staphylococ
18	229.5	13.3	953	16	Q99W48	Q99W48 staphylococ
19	210.5	12.2	947	2	Q86487	Q86487 staphylococ
20	204	11.8	1161	2	Q9X3M7	Q9X3M7 streptococ
21	204	11.8	1161	2	Q8RJ10	Q8RJ10 streptococ
22	201.5	11.6	1733	2	Q9K114	Q9K114 staphylococ
23	182	10.5	1039	2	P72534	P72534 streptococ
24	182	10.5	1160	2	Q8RM86	Q8RM86 streptococ
25	161.5	9.3	463	16	Q92DD5	Q92DD5 listeria in
26	157.5	9.1	462	16	Q8Y8L7	Q8Y8L7 listeria in
27	147.5	8.5	586	16	Q92FA3	Q92FA3 listeria in
28	145.5	8.4	2276	2	Q93TY6	Q93TY6 staphylococ
29	141	8.1	1386	16	Q92DL0	Q92DL0 listeria in
30	140.5	8.1	1166	12	Q8V7K1	Q8V7K1 chlorella v
31	140.5	8.1	2065	2	Q93DC7	Q93DC7 versinia en
32	137	7.9	439	16	Q8Y783	Q8Y783 listeria mo
33	137	7.9	4152	2	Q92HL3	Q92HL3 haemophilus
34	135	7.8	1983	16	Q9CJ55	Q9CJ55 lactococcus
35	134	7.7	1454	2	Q9F324	Q9F324 neisseria m
36	132.5	7.7	1457	16	Q9JXL6	Q9JXL6 neisseria m
37	132	7.6	1449	2	Q9F325	Q9F325 neisseria m
38	132	7.6	1449	16	Q5JWB4	Q5JWB4 neisseria m
39	128	7.4	1158	16	Q8YZU0	Q8YZU0 anabaena sp
40	128	7.4	1167	17	Q8TUJ9	Q8TUJ9 methanosarc
41	127	7.3	440	2	O07120	O07120 lactobacill
42	127	7.3	482	17	Q8TTC8	Q8TTC8 methanosarc
43	127	7.3	793	16	Q8YAG7	Q8YAG7 listeria mo
44	126.5	7.3	571	16	Q8YAG6	Q8YAG6 listeria mo
45	126.5	7.3	1029	2	Q52708	Q52708 rickettsia
46	126.5	7.3	1030	5	Q962B3	Q962B3 helicoverpa
47	125.5	7.2	1457	2	Q9X7H1	Q9X7H1 neisseria m
48	125.5	7.2	1464	12	Q8V7J9	Q8V7J9 chlorella v
49	125.5	7.2	2143	16	Q8RIP5	Q8RIP5 fusobacteri
50	124	7.2	456	17	Q8TIV8	Q8TIV8 methanosarc
51	123.5	7.1	1129	16	Q8XL11	Q8XL11 clostridium
52	123.5	7.1	1615	2	Q9F0P9	Q9F0P9 anabaena sp
53	123	7.1	1615	2	Q9F0P9	Q9F0P9 rickettsia
54	122.5	7.1	657	16	Q92DS2	Q92DS2 listeria in
55	122.5	7.1	1039	16	Q9ANV1	Q9ANV1 streptococ
56	122.5	7.1	1420	5	Q9NEV7	Q9NEV7 drosophila
57	122.5	7.1	1487	16	Q8YK40	Q8YK40 anabaena sp
58	122	7.0	445	17	Q8TLB5	Q8TLB5 methanosarc
59	122	7.0	1578	16	Q92E25	Q92E25 listeria in
60	121.5	7.0	1060	2	Q1LAR1	Q1LAR1 rickettsia
61	121.5	7.0	3890	16	Q99U53	Q99U53 staphylococ
62	121	7.0	1910	16	Q82BX3	Q82BX3 versinia pe
63	120.5	7.0	920	2	Q45664	Q45664 bacillus st
64	120.5	7.0	1268	2	Q9RER7	Q9RER7 bacillus sp
65	120	6.9	979	2	Q9XAS7	Q9XAS7 streptococ
66	120	6.9	1654	2	Q93QW9	Q93QW9 rickettsia
67	119.5	6.9	574	2	Q87360	Q87360 staphylococ
68	119.5	6.9	691	16	Q9RZS7	Q9RZS7 deinooccus
69	119.5	6.9	836	16	Q97GS5	Q97GS5 clostridium
70	119.5	6.9	1643	2	Q9F0P7	Q9F0P7 rickettsia
71	118.5	6.8	749	16	Q9CGA2	Q9CGA2 lactococcus
72	118.5	6.8	1124	17	Q9HKQ3	Q9HKQ3 thermoplasma
73	118.5	6.8	1369	12	Q89349	Q89349 paramesium
74	118	6.8	340	17	Q8TU51	Q8TU51 methanosarc
75	118	6.8	1873	2	Q924N7	Q924N7 enterococcu
76	118	6.8	1948	17	Q8T151	Q8T151 methanosarc
77	117.5	6.8	433	16	Q97TE6	Q97TE6 clostridium
78	117.5	6.8	1643	2	Q9F0P6	Q9F0P6 rickettsia
79	117.5	6.8	1758	16	Q9JMS5	Q9JMS5 escherichia
80	117.5	6.8	2747	2	Q9L800	Q9L800 aeromonas s
81	117	6.8	856	2	Q92FG9	Q92FG9 azotobacter
82	117	6.8	1713	3	Q8TGE1	Q8TGE1 saccharomyc
83	116.5	6.7	449	2	Q9S0W9	Q9S0W9 lactobacill
84	116.5	6.7	725	16	Q9CFY6	Q9CFY6 lactococcus
85	116.5	6.7	1441	16	Q9CFL1	Q9CFL1 lactococcus
86	116.5	6.7	3183	2	Q8RR93	Q8RR93 mycoplasma
87	116.5	6.7	4199	16	P74440	P74440 synechocyst
88	116	6.7	497	2	Q9VVG6	Q9VVG6 streptococ
89	116	6.7	657	2	Q93D79	Q93D79 bacillus th

90 116 6.7 784 16 Q8YAJ5 Q8YAJ5 listeria mo
 91 116 6.7 839 2 P77792 P77792 chlamydia p
 92 116 6.7 1570 5 Q9U0H8 Q9U0H8 plasmodium
 93 116 6.7 2340 5 O97298 O97298 plasmodium
 94 115.5 6.7 461 17 Q8TTC5 Q8TTC5 methanosarc
 95 115.5 6.7 1603 2 Q9KKA6 Q9KKA6 rickettsia
 96 115.5 6.7 2806 16 Q8RII19 Q8RII19 fusobacteri
 97 115.5 6.7 4936 16 Q8YKJ3 Q8YKJ3 anabaena sp
 98 115 6.6 785 16 Q92FC9 Q92FC9 listeria in
 99 115 6.6 858 5 P90545 P90545 entamoeba h
 100 115 6.6 1582 16 Q8Y9A5 Q8Y9A5 listeria mo

ALIGNMENTS

RESULT 1
 Q53653
 ID Q53653 PRELIMINARY; PRT; 933 AA.
 AC Q53653
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Clumping factor.
 DE Clumping factor.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94224142; PubMed=8170386;
 RA Cui L., Oquchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 aureus".
 RL Lancet 357:1225-1240(2001).
 DR EMBL; AP003360; BAB56973.1; .
 DR InterPro; IPR000515; BPD_transp.
 DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
 DR TIGRFAMS; TIGR01168; YSTRK_signal; 1.
 DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 933 AA; 97058 MW; EBS1A6DE2FF759F4 CRC64;
 Query Match 99.7%; Score 1727; DB 2; Length 933;
 Best Local Similarity 100.0%; Pred. No. 1.9e-87;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2 VAADAPAAAGTDITNLTNTVTVGIDSGTTVYPHQAGYVKLVNLYGFSVPNSAVKGDTEKTIYV 61
 221 VAADAPAAAGTDITNLTNTVTVGIDSGTTVYPHQAGYVKLVNLYGFSVPNSAVKGDTEKTIYV 280
 QY 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 121
 DB 281 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 340
 QY 122 PENVKKTGNVTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKNTNTYQTIYV 181
 DB 341 PENVKKTGNVTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKNTNTYQTIYV 400
 QY 182 NPSGDNVIAPVLTLGNLKPNTSDNALIDQONTSIKYKVDNAADLSSESFVNPENFEDVTN 241
 DB 401 NPSGDNVIAPVLTLGNLKPNTSDNALIDQONTSIKYKVDNAADLSSESFVNPENFEDVTN 460
 QY 242 SVNITFPNPQYKVFENTPDDQITTPYIVVNGHIDPNKSGDLALRSLTYGNSNIWRS 301
 DB 461 SVNITFPNPQYKVFENTPDDQITTPYIVVNGHIDPNKSGDLALRSLTYGNSNIWRS 520
 QY 302 MSWDNEVAFNNGSGSGDGDIDKPVVPEQDPE 331
 DB 521 MSWDNEVAFNNGSGSGDGDIDKPVVPEQDPE 550

RESULT 2
 Q932C5
 ID Q932C5 PRELIMINARY; PRT; 935 AA.
 AC Q932C5
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Fibrinogen-binding protein.
 DE FNB OR SAV0811.
 GN FNB OR SAV0811.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=158878;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oquchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 aureus".
 RL Lancet 357:1225-1240(2001).
 DR EMBL; AP003360; BAB56973.1; .
 DR InterPro; IPR000515; BPD_transp.
 DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
 DR TIGRFAMS; TIGR01168; YSTRK_signal; 1.
 DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 935 AA; 96950 MW; DC5A2D92CE3BA91C CRC64;
 Query Match 87.6%; Score 1517; DB 16; Length 935;
 Best Local Similarity 87.3%; Pred. No. 7.3e-76;
 Matches 288; Conservative 17; Mismatches 25; Indels 0; Gaps 0;
 QY 2 VAADAPAAAGTDITNLTNTVTVGIDSGTTVYPHQAGYVKLVNLYGFSVPNSAVKGDTEKTIYV 61
 DB 221 VAADAPAAAGTDITNLTNTVTVGIDSGTTVYPHQAGYVKLVNLYGFSVPNSAVKGDTEKTIYV 280
 QY 62 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 121
 DB 281 KELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYID 340
 QY 122 PENVKKTGNVTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKNTNTYQTIYV 181
 DB 341 PENVKKTGNVTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKNTNTYQTIYV 400
 QY 182 NPSGDNVIAPVLTLGNLKPNTSDNALIDQONTSIKYKVDNAADLSSESFVNPENFEDVTN 241
 DB 401 NPSGDNVIAPVLTLGNLKPNTSDNALIDQONTSIKYKVDNAADLSSESFVNPENFEDVTN 460
 QY 242 SVNITFPNPQYKVFENTPDDQITTPYIVVNGHIDPNKSGDLALRSLTYGNSNIWRS 301
 DB 461 QVRSIFPNANQYKVFENTPDDQITTPYIVVNGHIDPNKSGDLALRSLTYGNSNIWRS 520
 QY 302 MSWDNEVAFNNGSGSGDGDIDKPVVPEQDPE 331
 DB 521 MSWDNEVAFNNGSGSGDGDIDKPVVPEQDPE 550

RESULT 3
 Q99VJ4
 ID Q99VJ4 PRELIMINARY; PRT; 989 AA.
 AC Q99VJ4
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Fibrinogen-binding protein A, clumping factor.

GN CLFA OR SA0742.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21311952; PubMed-11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami A., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Tabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003131; BAB41975.1; -.
DR InterPro; IPR000515; BAP-transp.
DR InterPro; IPR001899; Gram_pos_anchor.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBR; UNKNOWN_1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;
Query Match 87.6%; Score 1517; DB 16; Length 989;
Best Local Similarity 87.3%; Pred. No. 7 8e-76;
Matches 288; Conservative 17; Mismatches 25; Indels 0; Gaps 0;
QY 2 VAADAPAGTDTITNLTNTVTVGIDSGTTVYPHOAGYVKLVNFGFSPNSAVKGDTEKITYP 61
Db 221 VAADAPAGTDTITNLTNTVTVGIDSGTTVYPHOAGYVKLVNFGFSPNSAVKGDTEKITYP 280
QY 62 KEINLVGVTSTAKVPPIMAGDOVLANGVIDSGNVIYTFDVTNKKDVKATLTMPAYID 121
Db 281 KEINLVGVTSTAKVPPIMAGDOVLANGVIDSGNVIYTFDVTNKKDVKATLTMPAYID 340
QY 122 PENVAKTGNVLTATIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTNNTYRTIYV 181
Db 341 PENVAKTGNVLTATIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTNNTYRTIYV 400
QY 182 NPSGDNVAPLVLTGNLKNPTDSNALIDQONTSIKVKYVDNAADLSSEYFVNPENFEDVTN 241
Db 401 NPSGDNVAPLVLTGNLKNPTDSNALIDQONTSIKVKYVDNAADLSSEYFVNPENFEDVTN 460
QY 242 SVNITFPNNQYKVFENTPDDQITTPYIVVNGHIDPNSKGLALRSTLYGYNSTINRS 301
Db 461 QVRISFPNNQYKVFENTPDDQITTPYIVVNGHIDPNSKGLALRSTLYGYNSTINRS 520
QY 302 MSWDNEVAFNNGSGDGDIDKVPVPEQDPE 331
Db 521 MSWDNEVAFNNGSGDGDIDKVPVPEQDPE 550
RESULT 4
Q93MH7 PRELIMINARY; PRT; 881 AA.
AC Q93MH7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Receptin Fbl precursor.
GN FBL.
OS Staphylococcus lugdunensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=28035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-2342;
RA Nilsson M., Ahlen J., Frykberg L., Guss B.;

RT "A fibrinogen-binding protein of Staphylococcus lugdunensis".
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF404823; AAK95649.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW SIGNAL.
FT CHAIN 1 39 POTENTIAL.
FT CHAIN 40 881 RECEPTIN_FBL.
SQ SEQUENCE 881 AA; 94251 MW; D4296C4959C4F19B CRC64;
Query Match 64.2%; Score 1112; DB 2; Length 881;
Best Local Similarity 62.6%; Pred. No. 1.5e-53;
Matches 201; Conservative 50; Mismatches 70; Indels 0; Gaps 0;
QY 11 TDTITNLTNTVTVGIDSGTTVYPHOAGYVKLVNFGFSPNSAVKGDTEKITYPKEINLVGVT 70
Db 214 SDISNKLSTVATIEAADIYPHKAEYVNLNRFQAPDDVQAGDSIKITIPQALNLTGVT 273
QY 71 STAKVPPIMAGDOVLANGVIDSGNVIYTFDVTNKKDVKATLTMPAYIDPENVKKTGN 130
Db 274 ATAKAPNIMAGDOVLATGIDEENLIYTFDVTNKNITGOISIPGIDPKNVTHTGK 333
QY 131 VTLATGIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTNNTYRTIYVNPSSGDNVIA 190
Db 334 VNLETSIGOTTAKKTVTVDEYKGFYNLSIKGTIDQIDKTNNTYRTIYVNPSSDVTVD 393
QY 191 PVLTKGNLKNPTDSNALIDQONTSIKVKYVDNAADLSSEYFVNPENFEDVTNSVNTIFPNP 250
Db 394 PYLRGGSIFGTSNVIIDEQNTSIKVKYKKAHLTDSYVVDPSNYEDVTSVKITTFPEK 453
QY 251 NOYKVFENTPDDQITTPYIVVNGHIDPNSKGLALRSTLYGYNSTINRSMSWDNEVAF 310
Db 454 GIYQINFNTEDQINSPPYVIVNGHIDPNSGLNLYLRSTLYGYDSNFTRVSMWDNEVEY 513
QY 311 NNGSGDGDIDKVPVPEQDPE 331
Db 514 HAGSGNGDIDKVPVPEQDPE 534
RESULT 5
Q99RD3 PRELIMINARY; PRT; 961 AA.
AC Q99RD3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fnb protein (Fibroectin-binding protein homolog).
GN FNB OR SAV2502 OR SA2290.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE-21311952; PubMed-11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami A., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003365; BAB58664.1; -.
DR EMBL; AF003137; BAB43593.1; -.
DR InterPro; IPR004237; Fb_bind.

```

DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR002965; P-rich_extensn.
DR Pfam: PF02986; Fn_bind; 1.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR PRINTS: PR01217; PRICHEXTENS.
DR TIGRFAMS: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 961 AA; 106010 MW; 364940F894EA4101 CRC64;

Query Match 24.0%; Score 416.5; DB 16; Length 961;
Best Local Similarity 29.1%; Pred. No. 3.7e-15;
Matches 101; Conservative 69; Mismatches 144; Indels 33; Gaps 9;

QY 2 VAADAPAACTDITNLTNTVTVGIDSGT-----SGTIVYVPHQAGYVKNLNGFSPVNSAVKGDTEK 57
DB 158 VAEAEKATGTDVTKNVEVETESLEGHNKDSNTVNPNAQRVTLKYKRFGEIGAKGDYFD 217

58 ITVPKELNLNGVTSTAKVPPIMAG--DOVLANGVIDSGNVIYTFDYVNTKDDVKATILT 115
218 FTLSNDNVETHGISTLRKVPKESSTEDKVMANGVINERTIRYTFDYVNTKDDVKATILT 277

116 MPAYIDPENVKKTGNVTATGISTTANKTVLDY-----EKYKGFYNLSIKGTIDQIDK 170
278 LNLFDPTTIVTQGSQKVEVTLGQNKVSKEFDIKYLDGVKDRMG-----VTVNGRIDTLNK 333

171 TNNYRTQIYVNPVSGDNVTAPVLTGKLPNTSDNALIDQNTSIKVKVDNAADLSSEYF 230
334 EGGKSFHFAVYKPNQSLTSVTQVTSYKQSA-----NNPTVKYKHIGSDLAESYI 389

231 V---NPENFEDVTSNVTNTPNPNOYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALR 287
390 AKLDDTSKFEDVTERVNSLTSNGGYTLNLGLDN--SKDYVIKYEYDQNAK--DLNFR 446

288 STLYGYSNINR-----SMSWDNEVAFNNGSGDGDIDKPVVP 326
447 THLSGYHYKYPYPPYPPYVQLTWNNGVAFYSNNAKGDGDKPNDP 493

*RESULT 6
Q53682 ID Q53682 PRELIMINARY: PRT: 940 AA.
AC Q53682;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibronectin binding protein B.
GN FNBB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8325-4;
RX MEDLINE=92111475; PubMed=1837266;
RA Jonsson K., Sigas C., Muller H.P., Lindberg M.;
RT "Two different genes encode fibronectin binding proteins in
RT Staphylococcus aureus. The complete nucleotide sequence and
RT characterization of the second gene.";
RL Eur. J. Biochem. 202:1041-1048(1991).
DR EMBL: X62992; CAA44726.1; .
DR InterPro: IPR004237; Fn_bind.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF02986; Fn_bind; 2.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 940 AA; 103555 MW; E35FBBCA507AE345 CRC64;

Query Match 23.6%; Score 409; DB 2; Length 940;

Best Local Similarity 28.9%; Pred. No. 9.3e-15;
Matches 101; Conservative 72; Mismatches 131; Indels 46; Gaps 11;

QY 2 VAADAPAACTDITNLTNTVTVGIDSGT-----TVPYVPHQAGYVKNLNGFSPVNSAV 51
DB 155 VVEETKATGTDVTKN-----VEEGSEIVGHKQDNTVNPNAERVTLYKKWKFGEGTK 209

52 KGDTEKITVPKELNLNGVTSTAKVPPIMAGD--QVLANGVIDSGNVIYTFDYVNTKDDV 110
210 AGDYFDFTLSNDNVETHGISTLRKVPKESSTEDKVMANGVINERTIRYTFDYVNTKDDV 269

111 KATLTPAYIDPENVKKTGNVTATGISTTANKTVLDY-----EKYKGFYNLSIKGTI 165
270 TAEISLNLFDITPTVTKGNQNVKVLGETTVYSKIFNIQYLGVRDNGW-----VTANGRI 325

166 DQIDKTNNTYRTIYVNPVSGDNVIAVPLVTGNL-----KPNVDSNALIDQNTSIKVKVDN 221
326 DTLNKVDGKFSHFAYMKPNQSLSSVTVTQVTKGNKPGVN-----NPTVKYKHIG 377

222 AADLSSEYFVNPN-----FEDVTSNVTNTPNPNOYKVEFNTPDQITTPYIVVNGHIDP 278
378 SDDLAEVYAKLDDVSKFEDVTDNMSLDFDTNNGYSLNFNLDQ--SKNYVIKYEYDYS 435

279 NSKGDLALRSTLYGYSNINRSMWDSNEVAFNNGSGDGDIDK---PVV 325
436 NA-SNLEFOTHLFGYNYNYTSLNTWKNGVAFYSNNAQGDGDKLKEPII 484

RESULT 7
Q9K113 ID Q9K113 PRELIMINARY: PRT: 931 AA.
AC Q9K113;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibronogen-binding protein SdrG.
GN SDRG.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K28;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RT epidermidis.";
RL Microbiology 146:1535-1546(2000).
DR EMBL: AF245042; AAF72510.1; .
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 931 AA; 102955 MW; 591E657D97027116 CRC64;

Query Match 20.4%; Score 353; DB 2; Length 931;
Best Local Similarity 30.6%; Pred. No. 1.1e-11;
Matches 101; Conservative 60; Mismatches 139; Indels 30; Gaps 14;

QY 13 ITNLTNTVTVGI--DSGTTVYVPHQAGYVKNLNGFSPVNSAVKGDTEKITVPKELNLNGVTS 71
DB 285 VTQDQ--SITEGYDDSDGIKAHAENLIYDVTFFVDDKVKSGDITVNLIDKKTIVPSDLTD 342

72 TAKVPPIMAGD--QVLANGVID--SGNVIYTFDYVNTKDDVKATLTPAYIDPENVKKTG 129
343 SFAIPKIDNSGEIATGTYDNTNQITVTFDYVDVKENIKAHKLTSYIDSKVPPNNN 402

130 ---NVTLATIGSTTANKTVLDYKFKFYNLSIKGTIDQIDKTNNTYRTIYVNPVSGD 186
403 TKLDVEYKTLSS--VNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYINPL-- 458

```

QY	187	NVIAPVL	TGNLKPNTDSNALIDQO	NTSIKVVKVQDNAA	DLSESYFV-NPENFEDV	TNSVNI	243
Db	459	RYSAKET	NVNISGNGDEG	STIIDSTIIKVVKVG	DNQLPDSNR	IYDSEYEDVTND	518
QY	246	TFPNPQY	KVEFNTPDQI	TPYTVVNGHIDP	NSKGD-----	LALRSTLYG	298
Db	519	QLGNNDV	NVNFG----	NIDSPYIKVISKY	DPN-KDYTTIQ	QVTVWQTINEY	571
QY	299	WRSMWDNE	VAFNNGSGG	IDKPVVPEQ	328		
Db	572	FRTASYD	NTIAFSTSSG	OGQ-DLP--	PEK	598	
*RESULT 8							
ID	070022	PRELIMINARY;		PRT;		1092	AA.
AC	070022;						
DT	01-AUG-1998	(TREMBLrel. 07, Created)					
DT	01-AUG-1998	(TREMBLrel. 07, Last sequence update)					
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)					
DT	Fibrinogen-binding protein precursor.						
OS	Staphylococcus epidermidis.						
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;						
OC	Staphylococcus.						
OC	NCBI_TaxID=1282;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=HB;						
RA	MEDLINE=98261511; PubMed=9596732;						
RA	Nilsson M., Frykberg L., Flock J.I., Pei L., Lindberg M., Guss B.;						
RT	"A Fibrinogen-binding protein of Staphylococcus epidermidis.";						
RL	Infect. Immun. 66:2666-2673(1998).						
DR	EMBL:	Y17116:	CAA76638.1:	-			
DR	InterPro:	IPR001899;	Gram_pos_anchor.				
DR	Pfam:	PF00746;	Gram_pos_anchor.				
DR	TIGRFAMS:	TIGR01167;	LPXTG_anchor;	1.			
DR	TIGRFAMS:	TIGR01168;	YSIRK_signal;	1.			
DR	PROSITE:	PS00343;	GRAM_POS_ANCHORING;	UNKNOWN_1.			
KW	Signal.						
FT	SIGNAL	1	51	POTENTIAL.			
FT	CHAIN	52	1092	FIBRINOGEN-BINDING PROTEIN.			
SO	SEQUENCE	1092	AA;	119292	MM;	6542BC39AAD8B984	CRC64;
Query Match							
Best Local Similarity 20.2%; Score 349; DB 2; Length 1092;							
Matches 100; Conservative 61; Mismatches 139; Indels 30; Gaps 14							
Db	13	ITNQLT	NVTVGI-DSGT	TVYPHQAGYVK	LYGFSVPNSAV	KDGTFTK	ITVPKELNLNGVTS
Db	287	VTDO--	SITEGYDD	SEGIKAHA	DAENLIYD	VTFEVDDK	VKSGDTMTVDIDKNTVP
Db	72	TAKVPP	IMAGD-QV	LANGVIDS-DGN	VLYT	TDYVNTK	DDVKATLTMPAYIDPENVK
QY	72	TAKVPP	IMAGD-QV	LANGVIDS-DGN	VLYT	TDYVNTK	DDVKATLTMPAYIDPENVK
Db	345	SFTPIK	NDNSGE	IIATGYDN	KNKQI	TYTFTDYV	DKYENIKAHUKLTSYIDKSKVP
QY	130	---NV	LATG	IGSTANKTV	LDVDEKYG	KFYNLSIK	GTIDQIDKNTNTYROT
Db	405	TKLDV	EYK	TALSS--	VNKTIT	VEYQRP	ENRTANLQSMFTNIDTNKH
QY	187	NVIAPVL	TGNLKPNTDS	NALIDQO	NTSIKVVKVQDN	AA	DLSESYFV-NPENFEDV
Db	461	RYSAKET	NVNISGNGDEG	STIIDSTIIKVVKVG	DNQLPDSNR	IYDSEYEDVTND	518
QY	246	TFPNPQY	KVEFNTPDQI	TPYTVVNGHIDP	NSKGD-----	LALRSTLYG	298
Db	521	QLGNNDV	NVNFG----	NIDSPYIKVISKY	DPN-KDYTTIQ	QVTVWQTINEY	571
QY	299	WRSMWDNE	VAFNNGSGG	IDKPVVPEQ	328		
Db	574	FRTASYD	NTIAFSTSSG	OGQ-DLP--	PEK	600	

[illegible]


```

SQ SEQUENCE 913 AA: 97247 MW: 38077C321F8D3E61 CRC64;
Query Match 18.0%; Score 312.5; DB 2; Length 913;
Best Local Similarity 24.9%; Pred. No. 1.9e-09;
Matches 92; Conservative 69; Mismatches 132; Indels 77; Gaps 15;

QY 6 APAAGTDTITNLTNTVVGIDSGTTVYPHOAGVYKLVNLYGFSVPNSAVKGDFTFKITVPKELN 65
DB ADAGKTNNVDKVTASNFLEK-TTDPNGSGNTFMAANFTVTDKVKSGDYF----- 257
QY 66 LNCVSTAKVPPIMAGDQVLANGVID-----SDGNVI----- 97
DB ADAGKTNNVDKVTASNFLEK-TTDPNGSGNTFMAANFTVTDKVKSGDYF----- 306
QY 98 YFTDVTNPKDVKATLTMPAIDPENVKKTGNVLTATIGISTTANKTVLVDYER----Y 153
DB 307 YFTDVTNPKDVKATLTMPAIDPENVKKTGNVLTATIGISTTANKTVLVDYER----Y 366
QY 154 GKFNLSIKGTIDQIDKTN--NTYRQTIYVNP-----GDNVIAPVLTGNLKPNTDSNALI 207
DB 367 DRKPGANISSQIIGVDTASGQNTYKQTVFVNPQKQVLTGNTWV--YIKGYQDKIESSGKV 424
QY 208 DOQNTSIKVKYVDNAADLSESVNP--ENFEDVNSV--NITFPNPQYKVEFNPDDQ 263
DB 425 SATDTKLRIEYVNDTSKLSDSYADPNDSNLKEVDQFKNRIYEHNPVASIKFG---D 480
QY 264 ITTPVIVVNGHIDPNSKGLALRSLTYGN---SNIIWRSMWDNEVAFNNGSGSDG 319
DB 481 IYKTVVLVEGHYDNTGKN---LKTQVTOENVDPVTNRDYSIFGWNENNVRYGGSGADG 537
QY 320 IDKPVVPEQP 329
DB 538 -DSAVNPKDP 546

RESULT 15
Q9KWX6 PRELIMINARY; PRT; 1171 AA.
AC Q9KWX6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Bone sialoprotein-binding protein.
GN BBP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=024;
RX MEDLINE=20115096; PubMed=10642520;
RA Tung H.S., Guss B., Hellman U., Persson L., Rubin K., Ryden C.;
RT "A bone sialoprotein-binding protein from Staphylococcus aureus: a
RT member of the staphylococcal Sdr family []."
RL Biochem. J. 345:611-619(2000).
DR EMBL; Y18653; CAB75732.1;
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRfams; TIGR01167; LPXTG_anchor; 1.
DR TIGRfams; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 1171 AA; 127123 MW; C5BC812F9DA5A884 CRC64;

Query Match 17.2%; Score 298; DB 2; Length 1171;
Best Local Similarity 29.2%; Pred. No. 1.6e-08;
Matches 93; Conservative 49; Mismatches 146; Indels 30; Gaps 10;

QY 2 VAADAPAGTDTITNLT-----NVTYVIGDGTTPVPHOAGVYKLVNLYGFSVPNSAVKGDFTFK 57
DB 264 VAQPAVASNNVNDLITVKQMITGKIDGVIQAHDEGHIIYTSDFKIDNAKAGADTWT 323
QY 58 IYVPEKLNNGVTSTAKVPPIMAGD---QVLANGVIDSGNVI-YTFTDVTNPKDVKAT 113
```

```

DB 324 VKYDKETIPSDITD--DFTPVDTDPGSEVIAKGFDLATKTIYKFTQVDRYENVNAK 381
QY 114 LMPAYIDPENVKKGTGNVTLATIGISTTANKTVLVDYKYGKFNLSIKGTIDQIDKTNW 173
DB 382 LELNSYIDKKEVPNETNLNLTATADKETSKNVKEVQKPIVKDESNIQSFSLDITTKH 441
QY 174 TYRQTIYVNP---SGDNVIAPVLTGNLKP-----TDSNALIDQOQNTSIKVKVDNAADL 225
DB 442 EVEQTIYVNPPLKLNKAKNTWVTRSGVADNGDYTGDSITID-SNTEIKVKVKSAGQQL 500
QY 226 SESYEV-NPENFEDVNTSNITFPNPQYKVEFNPDDQITTPYIVVNGHIDPNSKGLD 284
DB 501 POSNKIYDSQYEDVNTSVTI--NKNYGTNNANINFGDISAYIVKVVSKYTPGAEDDL 557
QY 285 ALR-----STLYGYNS 295
DB 558 AVQGVVRMTTNNKYNYS 575

RESULT 16
O86488 PRELIMINARY; PRT; 1315 AA.
AC O86488;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Sdr protein.
GN SDR.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=NEWMAN;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus."
RL Microbiology 144:3387-3395(1998).
DR EMBL; AJ005646; CAA06651.1;
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRfams; TIGR01167; LPXTG_anchor; 1.
DR TIGRfams; TIGR01168; YSIRK_signal; 1.
SQ SEQUENCE 1315 AA; 142775 MW; 58D4E1F48E6A689 CRC64;

Query Match 14.9%; Score 257.5; DB 2; Length 1315;
Best Local Similarity 26.9%; Pred. No. 3.2e-06;
Matches 88; Conservative 51; Mismatches 157; Indels 31; Gaps 12;

QY 13 ITNQLTNTVVGIDSGTTVYPHOAGVYKLVNLYGFSVPNSAVKGDFTFKITVPKELNAGVT-- 70
DB 251 ITSNTTLTVVDADKNNKIVPAQ--DYLSLKSQLITVDDKVKSGDYFTIKYSDTVQVYGLNPE 309
QY 71 -----STAKVPPIMAGDQVLANGVIDSGNVI-YTFTDVTNPKDVKATLTMPAIDPEN 124
DB 310 DIKNIGDIDKP---NNGETIATAKHDTANNLITYTFTDVTNPKDVKATLTMPAIDPEN 366
QY 125 VKKTGN-VTLATIGISTTANKTVLVDYKYGKFNLSIKG---TIDQIDKTNNT--YRQ 177
DB 367 IPVSKNDVEFNVTIGNTTTKTTANIQYDPYVNVNKNKSIKSAFTETVSHVGNENPGYKQ 426
QY 178 TIYVNPNSGDNVIAPVLTGNLKPNTDSNAL--IDQOQNTSIKVKVDNAADLSESVNPEN 235
DB 427 TIYVNPSENSLTNAKLKVAQVHSSYPNNIQLNKDQVTDIKIYQVPGVYLNKGYDVNTKE 486
QY 236 FEDVTNSV--NITFPNPQYKVEFNPDDQITTPYIVVNGHID-PNSKGLALRSLTYG 292
DB 487 LPTDVTNQLQKITGYGDNNSAVIDFGNAD----SAYVVMVNTKFTQNTSSEPTLVQMATLS 542
```



```

01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DR Ssrc protein.
GN SDRc.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEWMAN;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
DR EMBL: AJ005645; CAA06650.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal; 1.
PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN 1.
SQ SEQUENCE 947 AA; 10288 MW; 3C6EFD6E35121554 CRC64;

Query Match 12.2%; Score 210.5; DB 2; Length 947;
Best Local Similarity 23.1%; Pred. No. 0.00083;
Matches 83; Conservative 57; Mismatches 136; Indels 83; Gaps 16;

QY 6 APAAGTDITNQL--TNVTVGIDSG-----TTVPHOAGYVKLYGFSVPNSAVKGD 55
DB 178 APOGGTNDKVNKHFSNIDAKGHVNQTTGTTEFWATSDVLKLANITYDDSKVKG 237
QY 56 FKITVPKELNLNGVTSTAKVPPIM--AGDOVLANGVIDSDGN--VIYTFDYVNTKDV 113
DB 238 FTFKYQYRPGSVRLPSQTNLYNAQGNLIAGIYDSTNTTTFYTFYVQYTNVRS 297
QY 114 LTPMAYIDPEN--VKKTG---NVTLATGIGSTTANKTLVLDYKFKYKFNLSIKGT 168
DB 298 FEQVAFAKRNATTDKTAVMETL---GNDYSEEIIVDYG-----NKAQPLISST 347
QY 169 DKTNN---TYRQTIYVNPNGDNVIAPVLGTNL-----KNTDSNALIDQNTSIK 220
DB 348 NYINNEDLSRNTAYVNPQKNTYTKOTFTVNTLTGYKFNPN-----KNFKIYEV 397
QY 221 NAAIDSESYFVNPENEDVTNSNITFPNPN-----QYKV--EFNTPDDQ 263
DB 398 DQNOFVDSFTPTSKLKDVTDFDVIYSDNKTATVDMKGGTSSNKQYIIQOVAYPD 457
QY 264 ITTPYIVVYVNGHIDPNKSGDLALRSTLYGYNSNIWRSMWSDNEVAFNNGSGD 322
DB 458 STD-----NGKIDYTLDTD---KTKY-----SWNSYSNVNGSSTANGDQK 495

RESULT 20
QYX3M7
ID QYX3M7 PRELIMINARY; PRT; 1161 AA.
AC QYX3M7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Fibronectin-binding protein I.
GN PFBB.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A735;
RX MEDLINE=99270927; PubMed=10338474;
RA Rocha C.L., Fischetti V.A.;
RT "Identification and characterization of a novel fibronectin-binding
RT protein on the surface of group A streptococci.";

01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DR Ssrc protein.
GN SDRc.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEWMAN;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
DR EMBL: AJ005645; CAA06650.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal; 1.
PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN 1.
SQ SEQUENCE 947 AA; 10288 MW; 3C6EFD6E35121554 CRC64;

Query Match 11.8%; Score 204; DB 2; Length 1161;
Best Local Similarity 22.8%; Pred. No. 0.0024;
Matches 86; Conservative 72; Mismatches 145; Indels 74; Gaps 18;

QY 8 AAGTDITN--QLTNVTVGIDS--GTT-----VYPHOAGYVKLYGFSVPNSAV---K 52
DB 142 AGSKDVSSSLQLENPKMSVVYKGTESVSGAADFYRNHAAAFKMSFELKQDKSETINP 201
QY 53 GDTFKITVPKELNLNGVTSTAKVPPIM--AGDOVLANGVIDSDGN--VIYTFDYVNTK 109
DB 202 GDTFVLQDLDRRLNPKGISQ--DIPKLIYDSANSPLAIGKYHAENHOLIIVTFDYI 259
QY 110 VKATLTPAYIDPE-----NVKKT-----GNVTLATGIGSTTANKTLVLDY 151
DB 260 VOLSALSULENKEVLENTSISNFKSTIGGQETIKGTNVNLYGNESTKESNITNG 319
QY 152 KYGKFNLSIKGTIDIDKTNNTYROTIVYVNPNGDNVIAPVLGTNL-----KPN 207
DB 320 NVG-----GSIESYNTETGEFVYVYVNPNTNI--PYATMNLWGFGRAKRSNT 369
QY 208 DQNTS-----IKYKYVDNAADLSSEYFVNPENF---EDVTNSVNIITFPNPNQY 258
DB 370 NDANTSSAELGETQVYVEPEGEKLPSSYGVVDVTKLRTDITAGLNGFGOMTKRQ 429
QY 259 TDDQITTPYIVVYVNGHIDPNKSGDLALRSTLYGYNS-----NIIWRSMWSD 311
DB 430 --NNTQKAFIILKVTGKTQDQK--PLVWOSNLASFEGASEYAAFTVPVGNVYFQ 486
QY 312 NGSGSGDG---IDKPWV 325
DB 487 PSKSGSGSGKSEFTKPSI 503

RESULT 21
QYX3M7
ID QYX3M7 PRELIMINARY; PRT; 1161 AA.
AC QYX3M7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibronectin binding protein (Protein P2).
GN PFBI OR PRTF2.12.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A735;
RX MEDLINE=21843101; PubMed=11854196;
RA Bessen D.E., Kalia A.;
RT "Genomic localization of a T serotype locus to a recombinatorial zone
RT encoding extracellular matrix-binding proteins in Streptococcus
RT pyogenes.";
RL Infect. Immun. 70:1159-1167(2002).
DR EMBL: AY049088; AAL11469.1;
DR EMBL: AF447492; AAL86412.1;
SQ SEQUENCE 1161 AA; 127509 MW; BED4609D029B1FD7 CRC64;

Query Match 11.8%; Score 204; DB 2; Length 1161;
Best Local Similarity 22.8%; Pred. No. 0.0024;
Matches 86; Conservative 72; Mismatches 145; Indels 74; Gaps 18;

QY 8 AAGTDITN--QLTNVTVGIDS--GTT-----VYPHOAGYVKLYGFSVPNSAV---K 52
DB 142 AGSKDVSSSLQLENPKMSVVYKGTESVSGAADFYRNHAAAFKMSFELKQDKSETINP 201
QY 53 GDTFKITVPKELNLNGVTSTAKVPPIM--AGDOVLANGVIDSDGN--VIYTFDYVNTK 109
DB 202 GDTFVLQDLDRRLNPKGISQ--DIPKLIYDSANSPLAIGKYHAENHOLIIVTFDYI 259
QY 110 VKATLTPAYIDPE-----NVKKT-----GNVTLATGIGSTTANKTLVLDY 151
DB 260 VOLSALSULENKEVLENTSISNFKSTIGGQETIKGTNVNLYGNESTKESNITNG 319
QY 152 KYGKFNLSIKGTIDIDKTNNTYROTIVYVNPNGDNVIAPVLGTNL-----KPN 207
DB 320 NVG-----GSIESYNTETGEFVYVYVNPNTNI--PYATMNLWGFGRAKRSNT 369
QY 208 DQNTS-----IKYKYVDNAADLSSEYFVNPENF---EDVTNSVNIITFPNPNQY 258
DB 370 NDANTSSAELGETQVYVEPEGEKLPSSYGVVDVTKLRTDITAGLNGFGOMTKRQ 429
QY 259 TDDQITTPYIVVYVNGHIDPNKSGDLALRSTLYGYNS-----NIIWRSMWSD 311
DB 430 --NNTQKAFIILKVTGKTQDQK--PLVWOSNLASFEGASEYAAFTVPVGNVYFQ 486
QY 312 NGSGSGDG---IDKPWV 325
DB 487 PSKSGSGSGKSEFTKPSI 503
```


Db 370 GSGTSEFTKP 379

RESULT 24

Q8RM86

ID Q8RM86 PRELIMINARY; PRT; 1160 AA.
AC Q8RM86;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibronectin binding protein.
GN FBPI.

OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.

NCBI_TaxID=1314;

[1]

SEQUENCE FROM N.A.

STRAIN=B737;

MEDLINE=21843101; PubMed=11854196;

Bessen D.E., Kalia A.;

"Genomic localization of a T serotype locus to a recombinatorial zone
encoding extracellular matrix-binding proteins in Streptococcus
pyogenes.";

RL Infect. Immun. 70:1159-1167(2002).

DR EMBL; AY049089; AAL11472.1; -.

SQ SEQUENCE 1160 AA; 128148 MW; 097C1726A5126133 CRC64;

Query Match

Best Local Similarity 10.5%; Score 182; DB 2; Length 1160;

Matches 81; Conservative 78; Mismatches 149; Indels 62; Gaps 18;

QY 8 AAGTDITN-QLTNTVTGI-----DSGTTVPHOAGYVKLVNFGFSPNSAV---K 52

Db 142 AGSKDVSSSLQLENPKMSVSKYGEQKTSNSADFYRNAHYFKMSFELKQDKSETINP 201

QY 53 GDTFKITVPKELNLNGVTSTAKVPPIM--AGDOVLANGVIDSDGN-VIYTTDYVNTKDD 109

Db 202 GDTFVLQDLRLNPKGISQ--DIPKIIYDSENSPLAICKYDAKTHQLTYFTNTYAGLDK 259

QY 110 VKATLTMPAYIDPENVKKTGNVT-LATGIG--STTANKTVLVLDY----EKYKGFVN---L 159

Db 260 VLSAELSLFLENKEVLNTNISDFKSTIGGQETIKYKTVNVLVGNSTKESNYITNCLS 319

QY 160 SIKGITDIDKNTNTYRTIYVNPSPGDNVIAPLVTGLN-----KPTDSNALIDQON 211

Db 320 NVGSIESTYNTETGEFVYVTVNPNRNPYAVL--NLWGFAKRTAOGENDNSLVSSAQL 377

QY 212 TSIKYKYVDNAADLSESYFVNPNFEDVTVNSVNITFPNP-----QYKVEFNTPDQIT 265

Db 378 TGYDIYEPHNYRLPTSYGVVDISRL-NLRKDLKALPQSGTQOGANKRLRIDFG--ENLQG 434

QY 266 TPYIVVYNGHIDPNKSGDLALRSTLYGYSNIIWRS-----MSWDNEVAFNNGSGS 316

Db 435 KAFVVKYVTKGRADQSGK-ELIVQSHLSFNN---MGSYKTLRPNSHVSFTNEIALSPSKGS 490

QY 317 GDGIDKPVVP 326

Db 491 GSGTSEFTKP 500

RESULT 25

Q92DD5

ID Q92DD5 PRELIMINARY; PRT; 463 AA.

AC Q92DD5;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Hypothetical protein lin0879.

GN Lin0879.

OS Listeria innocua.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Listeriaceae; Listeria.

OX NCBI_TaxID=1642;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CLIP 11262 / SEROVAR 6A;

RX PubMed=11679669;

RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetoui F., Couve E., de Daruvar A., Dehous P.,

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordisiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,

RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;

RT "Comparative genomics of *Listeria species*.";

RL Science 294:849-852(2001).

DR EMBL; AL596166; CAC96111.1; -.

DR ListList; LIN00879;

DR InterPro; IPR001899; Gram_pos_anchor.

DR InterPro; IPR002482; LysM.

DR Pfam; PF01476; LysM; 1.

DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 463 AA; 49939 MW; D1C15FB01E33C07F CRC64;

Query Match

Best Local Similarity 24.9%; Score 161.5; DB 16; Length 463;

Matches 88; Conservative 48; Mismatches 129; Indels 89; Gaps 18;

QY 9 AGTDI-TNQLTNTVTGIDSGTTVP-HOAGYVKLVNFGFSPNSAVKGTGFKITVPKELNL 66

Db 25 AATDYGSSFFTNVALQNGQDPVTFKNSKVRVAYDFVITQPVVSGTWTLTITPDQLK 84

QY 67 NGVTSTAKVPPIMAGDQVLANGVID-SDGNVIYTFDYVNTKDDVKATLTMPAYIDPEN 125

Db 85 ---INFGFPVNDASGNTIANATIDPATGTTTLTFHYVNTHTNLSGSLFYFNATFNSKNI 141

QY 126 KK-----TGNTLATGIGSTTANKTVLVLDYKYGKGFYNSIKGTI 165

Db 142 QTDQVNPQFPVNNVTQINTYISKVTGSGTGSPIT-----VFKQGRM 185

QY 166 DQIDKTNNTYRTIYVNPSPGDNVIAP-----VLTGNLKPNTDSNALIDQNTSIKVKYVDN 221

Db 186 D--DKDTSIIHWTTLN---NALTPIDNAIYDTL--GTQNLN---GNATIK-YRDAN 233

QY 222 AADLSESYFVNP-----ENFEDVTVNSVNIITFPNPQYKVEFNTPDQITTPYIVVNGH 275

Db 234 KKVLSLN--VQPIALDSNRNPFELTIGTLN-----NQSV--ITYDKITTKQKSYTN-- 281

QY 276 IDPNKSGDLALRSTLYGYSNIIWRSKWDNEVAFNNGSGSGDIDKPVVPEQP 329

Db 282 -----KATLSGDNLDANSRNATVND--YGGGGQGTGTPPTPPVKEEP 322

Search completed: June 23, 2003, 14:03:29

Job time : 48.8422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

June 23, 2003, 13:59:31 ; Search time 8.15248 Seconds
(without alignments)
1426.837 Million cell updates

Title: US-10-056-052A-20

Perfect score:

Sequence: 1 QVOLKESGPGLVKPSQTL SI.....YYGYDGFVYWGCGTLVTSS 121

Scoring table: BLOSUM62

scoring cable: BLOSUM02
Gapop 10.0 , Gapext 0.5

atched: 283224 seqs. 96134422 residues

number of hits satisfying chosen parameters. 283224

Minimum DB seq length: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 100 summaries

Database :

```
Database :
PIR_3: *
1: pir1: *
```

```
1: pir1: *
2: pir2: *
```

```
2: p1r2: *
3: p1r3: *
```

3: p1r3:*

4: p1r4:*

7. 2114.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	534.5	82.6	117	2	S38563	Ig heavy chain v r
2	489	75.6	114	2	S11106	Ig heavy chain v r
3	489	75.6	140	2	S55028	Ig heavy chain v r
4	488	75.4	127	2	B31807	Ig heavy chain v r
5	487.5	75.3	122	2	S20809	Ig heavy chain v r
6	486.5	75.2	139	2	A32456	Ig heavy chain pre
7	482.5	74.6	122	2	A45049	Ig heavy chain v r
8	482	74.5	144	1	G2MS14	Ig heavy chain pre
9	478.5	74.0	135	2	S31913	Ig gamma-2a chain
10	477	73.7	112	2	S11100	Ig heavy chain v r
11	477	73.7	113	2	S11101	Ig heavy chain v r
12	476.5	73.6	115	2	S26470	Ig heavy chain v r
13	472	73.0	116	2	S11102	Ig heavy chain v r
14	470	72.6	140	2	S14238	Ig gamma-1 chain p
15	466.5	72.1	116	2	S42484	Ig heavy chain v r
16	465.5	71.9	141	2	S52446	Ig heavy chain v r
17	465.5	71.9	144	2	S11244	Ig gamma-2a chain
18	465	71.9	121	2	D30560	Ig heavy chain v r
19	464.5	71.8	120	2	PL0087	Ig heavy chain v r
20	464	71.7	114	2	S11099	Ig heavy chain v r
21	463	71.6	111	2	S26324	Ig heavy chain v r
22	459	70.9	115	2	S11103	Ig heavy chain v r
23	457.5	70.7	118	2	S32786	Ig heavy chain (an
24	457	70.6	115	2	S11107	Ig heavy chain v r
25	452	69.9	110	2	S26323	Ig heavy chain v r
26	452	69.9	118	2	PQ0266	Ig heavy chain v r
27	451	69.7	112	2	S11108	Ig heavy chain v r
28	450	69.6	231	2	PC3155	Ig gamma-2b chain
29	448.5	69.3	107	2	S14492	Ig heavy chain v r

ALIGNMENTS

RESULT 1

S38563 Ig heavy chain V region (ASWS1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S38563
R:Monestier, M.; Losman, L.J.; Novick, K.E.; Aris, J.P.
submitted to the EMBL Data Library, September 1993
A:Description: Molecular analysis of mercury-induced anti-nucleolar antibodies in H-2s M
A:Reference number: S38559
C:Keywords: heterotetramer; immunoglobulin
F;34-116/Domain: immunoglobulin homology <IMM>
Query Match 82.6%; Score 534.5; DB 2; Length 117;
Best Local Similarity 85.0%; Pred. No. 9.3e-43;
Matches 102; Conservative 5; Mismatches 10; Indels 3; Gaps 1;
Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGIMWGNGNDYN 60
Db 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGIMWGNGNDYN 60
Qy 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKEGFYGYDGFVYWGQGLTVTS 120
Db 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKEGFYGYDGFVYWGQGLTVTS 117

RESULT 2

S11106 Ig heavy chain V region (clone NQ5-96.2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 19-Mar-1997 #sequence_revision 23-Aug-1997 #text_change 21-Jan-2000
C:Accession: S11106
R:Kaartinen, M.; Griffiths, G.M.; Markham, A.F.; Milstein, C.
Nature 304, 320-324, 1983
A:Title: mRNA sequences define an unusually restricted IgG response to 2-phenyloxazalone
A:Reference number: S07331; MUID:83271467; PMID:6877353
A:Accession: S11106
A:Molecule type: mRNA
A:Residues: 1-114 <NAT>
Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F;15-97/Domain: immunoglobulin homology <IMM>
Query Match 75.6%; Score 489; DB 2; Length 114;
Best Local Similarity 76.9%; Pred. No. 1.5e-38;
Matches 93; Conservative 8; Mismatches 12; Indels 8; Gaps 1;
Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGIMWGNGNDYN 60
Db 1 QVQLKESGPGLVKPSQTLSTICTVSGASLSISYGVHVRQPPGKGLWLGIMWGNGNDYN 60
Qy 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKEGFYGYDGFVYWGQGLTVTS 120
Db 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKEGFYGYDGFVYWGQGLTVTS 112

RESULT 3

S55028 Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 23-Jul-1999

C:Accession: S55028

R:Jeffrey, P.D.; Schildbach, J.F.; Chang, C.Y.; Kussie, P.H.; Margolies, M.N.; Sherif
J. Mol. Biol. 248, 344-360, 1995
A:Title: Structure and specificity of the anti-digoxin antibody 40-50.
A:Reference number: S55027; MUID:95257394; PMID:7739045

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-140 <JEF>
A:Cross-references: EMBL:L31403; NID:9476717; PID:AAA38191.1; PID:9476718
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F;34-116/Domain: immunoglobulin homology <IMM>

Query Match 75.6%; Score 489; DB 2; Length 140;
Best Local Similarity 77.9%; Pred. No. 1.9e-38;
Matches 95; Conservative 8; Mismatches 17; Indels 2; Gaps 2;
Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGIMWGNGNDYN 60
Db 20 QVHLKESGPGLVKPSQTLSTICTVSGFSLTYGVHVRQPPGKGLWLGIMWGNGNDYN 79
Qy 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCAR-KGEFFYYGDFVYWGQGLTVTV 119
Db 80 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARFEASYDY-AVDYWGQGLTVTV 138
Qy 120 SS 121
Db 139 SS 140

RESULT 4

B31807 Ig heavy chain V region (PAC1) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Jul-1989 #sequence_revision 20-Jul-1989 #text_change 23-May-1997
C:Accession: B31807
R:Taub, R.; Gould, R.J.; Garsky, V.M.; Ciccarone, T.M.; Hoxie, J.; Friedman, P.A.; Sh
J. Biol. Chem. 264, 259-265, 1989
A:Title: A monoclonal antibody against the platelet fibrinogen receptor contains a se
A:Reference number: A31807; MUID:89079661; PMID:2909518
A:Accession: B31807
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-127 <TAU>
Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F;15-97/Domain: immunoglobulin homology <IMM>

Query Match 75.4%; Score 488; DB 2; Length 127;
Best Local Similarity 72.4%; Pred. No. 2.1e-38;
Matches 92; Conservative 11; Mismatches 18; Indels 6; Gaps 1;
Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGIMWGNGNDYN 60
Db 1 QVQLKESGPGLVKPSQTLSTICTVSGFSLTYGVHVRQPPGKGLWLGIMWGNGNDYN 60
Qy 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKEGFYGYDGFVYWGQGLTV 114
Db 61 AAFISRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKEGFYGYDGFVYWGQGLTV 120
Qy 115 TLTVSS 121
Db 121 TSVTVSS 127

RESULT 5

S20809 Ig heavy chain V region (hybridoma C8) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C:Accession: S20809
R:Hoogenboom, H.; Dubois, P.; Raus, J.; Volckaert, G.

Ig gamma-2A chain precursor - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 23-Jul-1999
C:Accession: S31913
R:Bespalov, I.A.; Hiyonov, P.A.; Lukashevich, L.V.; Lunev, V.E.; Tribush, S.S.; Gaponova
submitted to the EMBL data Library, January 1993
A:Reference number: S31913
A:Accession: S31913
A:Molecule type: mRNA
A:Residues: 1-135 <BES>
A:Cross-references: EMBL:X70822; NID:g57921; PIDN:CAA50153.1; PID:g57922
A:Experimental source: strain BAUB/c
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-11/DNA: signal sequence #status predicted <SIG>
F:12-135/Product: Ig gamma-2A chain (fragment) #status predicted <MAT>
F:34-116/DNA: immunoglobulin homology <IMM>

Query Match 74.0%; Score 478.5; DB 2; Length 135;
Best Local Similarity 74.4%; Pred. No. 1.7e-37;
Matches 90; Conservative 13; Mismatches 13; Indels 5; Gaps 1;

1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGIMGGNTDYN 60
20 QVQLKESGPGLVKPSQTLSTICTVSGFSLSRYSVHVRQPPGKGLWLGIMGGNTDYN 79
61 SALKSRSLTSKDNKSNQVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVYWGQGLTVTVS 120
80 AAFISRLSISKDNKSNQVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVYWGQGLTVTVS 134
121 S 121
135 S 135

RESULT 10
S11100
Ig heavy chain V region (clone N02-20.5.3) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 23-Aug-1997 #text_change 21-Jan-2000
C:Accession: S11100
R:Kaartinen, M.; Griffiths, G.M.; Markham, A.F.; Milstein, C.
Nature 304, 320-324, 1983
A:Title: mRNA sequences define an unusually restricted IgG response to 2-phenylloxazolone
A:Reference number: S07331; MUID:83271467; PMID:6877353
A:Accession: S11100
A:Molecule type: mRNA
A:Residues: 1-112 <NAT>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:15-97/DNA: immunoglobulin homology <IMM>

Query Match 73.7%; Score 477; DB 2; Length 112;
Best Local Similarity 76.7%; Pred. No. 2e-37;
Matches 92; Conservative 8; Mismatches 12; Indels 8; Gaps 1;

1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGIMGGNTDYN 60
20 QVQLKESGPGLVKPSQTLSTICTVSGFSLSRYSVHVRQPPGKGLWLGIMGGNTDYN 60
61 SALKSRSLTSKDNKSNQVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVYWGQGLTVTVS 120
80 AAFISRLSISKDNKSNQVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVYWGQGLTVTVS 112

RESULT 11
S11101
Ig heavy chain V region (clone N02-48.2.2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 21-Jan-2000
C:Accession: S11101
R:Kaartinen, M.; Griffiths, G.M.; Markham, A.F.; Milstein, C.
Nature 304, 320-324, 1983
A:Title: mRNA Sequences define an unusually restricted IgG response to 2-phenylloxazolone

A:Reference number: S07331; MUID:83271467; PMID:6877353
A:Accession: S11101
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <KAA>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:15-97/DNA: immunoglobulin homology <IMM>

Query Match 73.7%; Score 477; DB 2; Length 113;
Best Local Similarity 76.0%; Pred. No. 2e-37;
Matches 92; Conservative 9; Mismatches 12; Indels 8; Gaps 1;

1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGIMGGNTDYN 60
1 QVQLKESGPGLVKPSQTLSTICTVSGFSLSRYSVHVRQPPGKGLWLGIMGGNTDYN 60
61 SALKSRSLTSKDNKSNQVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVYWGQGLTVTVS 120
61 SALKSRSLTSKDNKSNQVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVYWGQGLTVTVS 112
121 S 121
113 A 113

RESULT 12

S26470
Ig heavy chain V region - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S26470
R:Kavaler, J.
submitted to the EMBL Data Library, April 1991
A:Reference number: S26459
A:Accession: S26470
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-115 <KAV>
A:Cross-references: EMBL:X59109; NID:g51948; PIDN:CAA41835.1; PID:g51949
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-97/DNA: immunoglobulin homology <IMM>

Query Match 73.6%; Score 476.5; DB 2; Length 115;
Best Local Similarity 76.9%; Pred. No. 2.2e-37;
Matches 90; Conservative 10; Mismatches 14; Indels 3; Gaps 2;

1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGIMGGNTDYN 60
1 QVQLKESGPGLVKPSQTLSTICTVSGFSLSRYSVHVRQPPGKGLWLGIMGGNTDYN 60
61 SALKSRSLTSKDNKSNQVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVYWGQGLTVTVS 116
61 AAFISRLSISKDNKSNQVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVYWGQGLTVTVS 115

RESULT 13

S11102
Ig heavy chain V region (clone N05-4.3.1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 23-Aug-1997 #text_change 21-Jan-2000
C:Accession: S11102
R:Kaartinen, M.; Griffiths, G.M.; Markham, A.F.; Milstein, C.
Nature 304, 320-324, 1983
A:Title: mRNA sequences define an unusually restricted IgG response to 2-phenylloxazolone
A:Reference number: S07331; MUID:83271467; PMID:6877353
A:Accession: S11102
A:Molecule type: mRNA
A:Residues: 1-116 <NAT>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:15-97/DNA: immunoglobulin homology <IMM>

Query Match 73.0%; Score 472; DB 2; Length 116;

Query Match 72.1%; Score 466.5; DB 2; Length 116;
Best Local Similarity 72.7%; Pred. No. 1.9e-36;
Matches 88; Conservative 13; Mismatches 15; Indels 5; Gaps 1;

```

RESULT 17
S11244
I9 gamma-2a chain precursor - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 16-Aug-1996
C:Accession: S11244
R:Wellman, A.A.; Meares, C.F.
Nucleic Acids Res. 18, 5281, 1990
A:Title: Sequences of the Lym-1 antibody heavy and light chain variable regions.
A:Reference number: S11244; MUID:90384832; PMID:2119497
A:Accession: S11244
A:Molecule type: mRNA
A:Residues: 1-144 <WEL>
A:Cross-references: EMBL:X53483
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin
F:34-116/Domain: immunoglobulin homology <IMM>

Query Match      71.9%; Score 465.5; DB 2: Length 144;
Best Local Similarity 75.4%; Pred. No. 3e-36;
Matches 92; Conservative 8; Mismatches 17; Indels 5; Gaps 2;

```



```
. QY 69 ISKDNKNOVFLKMNLSLTADTAIVYCARKEFYFG-YDGFVWGGTGLTVSS 121
      |||||:|||||: ||| ||||| ||| ||| ||||| |||||
Db 61 ISKDNKSOVFLKMNLSLTDDTARYCAR---VYGNVEAMDYWGQTSVTSS 111

RESULT 22
S11103
Ig heavy chain V region (clone NQ5-61.1.2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 23-Aug-1997 #text_change 21-Jan-2000
C:Accession: S11103
R:Kaartinen, M.; Griffiths, G.M.; Markham, A.F.; Milstein, C.
Nature 304, 320-324, 1983
A:Title: mRNA sequences define an unusually restricted IgG response to 2-phenyloxazolol
A:Reference number: S07331; MUID:83271467; PMID:6877353
A:Accession: S11103
A:Molecule type: mRNA
A:Residues: 1-115 <NAT>
F:15-97/Domain: immunoglobulin V region; immunoglobulin homology <IMM>

Query Match 70.9%; Score 459; DB 2; Length 115;
Best Local Similarity 73.6%; Pred. No. 9.5e-36;
Matches 89; Conservative 9; Mismatches 15; Indels 8; Gaps 1;

QY 1 QVQLKESGPGLVKPSQTLSTICTSGFSLRSYSVHVVRRPPGKGLWGLMIGWGNTDYN 60
      |||||:|||||: ||| ||||| ||| ||| ||||| |||||
Db 1 QVQLKESGPGLVAPSQSLSTICTVSGXSLTSGVHVVRRPPGKGLWGLVWAGGSTNYN 60

QY 61 SALKSRLSISKDNKNOVFLKMNLSLTADTAIVYCARKEFYFGYDGFVWGGTGLTVSS 120
      ||| |||||:|||||: ||||| ||||| ||||| |||||
Db 61 SALKSRLSISKDNKSKXVFLKMNLSLTDDTARYCAR-----DRGAYWGXTGLTVSS 112

QY 121 S 121
Db 113 A 113

RESULT 23
S32786
Ig heavy chain (anti-biotin) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S32786
R:Bagci, H.; Kohen, F.; Kuscuoğlu, U.; Bayer, E.A.; Wilchek, M.
Proc. Natl. Acad. Sci. USA 90, 11993-11997, 1993
A:Title: Monoclonal anti-biotin antibodies simulate avidin in the recognition of biotin.
A:Reference number: S32786; MUID:93245957; PMID:8482366
A:Accession: S32786
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-118 <BAG>
A:Cross-references: G8:S59639; NID:q299965; PIDN:AA26438.1; PID:q299966
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 70.7%; Score 457.5; DB 2; Length 118;
Best Local Similarity 72.1%; Pred. No. 1.3e-35;
Matches 88; Conservative 11; Mismatches 18; Indels 5; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTICTSGFSLRSYSVHVVRRPPGKGLWGLMIGWGNTDYN 60
      |||||:|||||: ||| ||||| ||| ||| ||||| |||||
Db 1 QVQLKESGPGLVAPSQSLSTICTVSGFSLTAYGVVRRPPGKGLWGLVWAGGSTNYN 60

QY 61 SALKSRLSISKDNKNOVFLKMNLSLTADTAIVYCARKEFYFGYD-GFYVWGGTGLTV 119
      ||| |||||:|||||: ||||| ||||| ||||| |||||
Db 61 SGLMSRLSIRKDNKSKQVFLTMNSLTDDTARYCYVK----HTNWDGGFAYWGQTTVT 116

QY 120 SS 121
Db 117 SS 118
```

RESULT 24

S11107

Ig heavy chain V region (clone NQ2-45.10.4) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 21-Jan-2000

C:Accession: S11107

R:Kaartinen, M.; Griffiths, G.M.; Markham, A.F.; Milstein, C.

Nature 304, 320-324, 1983

A:Title: mRNA sequences define an unusually restricted IgG response to 2-phenyloxazolol

A:Reference number: S07331; MUID:83271467; PMID:6877353

A:Accession: S11107

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-115 <RAA>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 70.6%; Score 457; DB 2; Length 115;

Best Local Similarity 71.9%; Pred. No. 1.5e-35;

Matches 87; Conservative 11; Mismatches 15; Indels 8; Gaps 1;

QY 1 QVQLKESGPGLVKPSQTLSTICTSGFSLRSYSVHVVRRPPGKGLWGLMIGWGNTDYN 60

Db 1 QVQLKQKQPGXVQXSQSSITCTVSGFSLTSYGVHVVRRPPGKGLWGLVWISGGSTDYN 60

QY 61 SALKSRLSISKDNKNOVFLKMNLSLTADTAIVYCARKEFYFGYDGFVWGGTGLTVSS 120

Db 61 AAFISRLSISKDNKSKQVFFKMNLSLTADTAIVYCAR-----NGDYGWGGTSTVTS 112

QY 121 S 121

Db 113 S 113

RESULT 25

S26323

Ig heavy chain V region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 20-Jun-2000

C:Accession: S26323

R:Stark, S.E.; Caton, A.J.

J. Exp. Med. 174, 613-624, 1991

A:Title: Antibodies that are specific for a single amino acid interchange in a protein

A:Reference number: S26309; MUID:91341421; PMID:1908510

A:Accession: S26323

A:Molecule type: mRNA

A:Residues: 1-110 <STA>

A:Cross-references: EMBL:X59184; NID:g52061; PIDN:CAA41894.1; PID:g1334033

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:6-87/Domain: immunoglobulin homology <IMM>

Query Match 69.9%; Score 452; DB 2; Length 110;

Best Local Similarity 78.6%; Pred. No. 4e-35;

Matches 88; Conservative 6; Mismatches 16; Indels 2; Gaps 1;

QY 10 GLVPSQTLSTICTSGFSLRSYSVHVVRRPPGKGLWGLMIGWGNTDYNLSALKSRLSI 69

Db 1 GLVPSQSLSTICTVSGFSLTGYGVNVRPPGKGLWGLMIGWGNTDYNLSALKSRLSI 60

QY 70 SKDNKNOVFLKMNLSLTADTAIVYCARKEFYFGYDGFVWGGTGLTVSS 121

Db 61 SKDNKSOVFLKMNLSLTDDTARYCARVP--YGNVDMDYWGQTSVTYSS 110

Search completed: June 23, 2003, 14:04:15

Job time : 9.15248 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 4.50532 Seconds

(without alignments)
1113.936 Million cell updates/sec

Title: US-10-056-052a-20

Perfect score: 647

Sequence: 1 QVQLKESGPGCLVKPSQTLSTL.....YYGYDGFVYWGQGLTVTVSS 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	482	74.5	144	1	HV43_MOUSE
2	434	67.1	115	1	HV44_MOUSE
3	423	65.4	116	1	HV45_MOUSE
4	396.5	61.3	135	1	HV02_XENLA
5	374.5	57.9	122	1	HV3G_HUMAN
6	370.5	57.3	146	1	HV2I_HUMAN
7	365	56.4	137	1	HV46_MOUSE
8	349.5	54.0	122	1	HV3H_HUMAN
9	349	53.9	121	1	HV3J_HUMAN
10	346	53.5	117	1	HV2G_HUMAN
11	344	53.2	136	1	HV01_XENLA
12	342	52.9	119	1	HV2C_HUMAN
13	338.5	52.3	116	1	HV61_MOUSE
14	338	52.2	129	1	HV2F_HUMAN
15	335.5	51.9	126	1	HV3K_HUMAN
16	333	51.5	113	1	HV47_MOUSE
17	333	51.5	119	1	HV3I_HUMAN
18	332.5	51.4	119	1	HV40_MOUSE
19	329.5	50.9	136	1	HV16_MOUSE
20	328.5	50.8	114	1	HV37_MOUSE
21	325.5	50.3	119	1	HV3B_HUMAN
22	325.5	50.3	126	1	HV2A_HUMAN
23	324.5	50.2	119	1	HV38_MOUSE
24	322	49.8	118	1	HV39_MOUSE
25	322	49.8	142	1	HV01_RAT
26	318.5	49.2	116	1	HV60_MOUSE
27	317.5	49.1	120	1	HV2B_HUMAN
28	316.5	48.9	116	1	HV3T_HUMAN
29	316	48.8	115	1	HV3D_HUMAN
30	316	48.8	123	1	HV25_MOUSE
31	314.5	48.6	122	1	HV3A_HUMAN
32	314	48.5	117	1	HV17_MOUSE
33	312	48.2	125	1	HV2D_HUMAN
					P01817 homo sapien

34	311	48.1	116	1	HV05_CARAU
35	310.5	48.0	120	1	HV3U_HUMAN
36	310	47.9	119	1	HV3L_HUMAN
37	309	47.8	115	1	HV32_MOUSE
38	308.5	47.7	114	1	HV2A_RABIT
39	308	47.6	117	1	HV62_MOUSE
40	308	47.6	120	1	HV3E_HUMAN
41	307.5	47.5	139	1	HV07_MOUSE
42	305	47.1	117	1	HV2B_RABIT
43	304	47.0	113	1	HV30_MOUSE
44	303.5	46.9	122	1	HV20_MOUSE
45	303.5	46.9	136	1	HV2C_RABIT
46	303.5	46.9	147	1	HV2H_HUMAN
47	303	46.8	115	1	HV3F_HUMAN
48	302.5	46.8	117	1	HV3C_HUMAN
49	301	46.5	117	1	HV02_CANFA
50	301	46.5	123	1	HV18_MOUSE
51	301	46.5	123	1	HV19_MOUSE
52	300	46.4	115	1	HV33_MOUSE
53	300	46.4	138	1	HV48_MOUSE
54	299	46.2	113	1	HV27_MOUSE
55	299	46.2	117	1	HV41_MOUSE
56	299	46.2	123	1	HV22_MOUSE
57	298.5	46.1	118	1	HV51_MOUSE
58	298	46.1	113	1	HV29_MOUSE
59	296	45.7	113	1	HV31_MOUSE
60	295	45.6	121	1	HV2E_HUMAN
61	295	45.6	123	1	HV23_MOUSE
62	293.5	45.4	119	1	HV34_HUMAN
63	293.5	45.4	122	1	HV21_MOUSE
64	293	45.3	113	1	HV28_MOUSE
65	292	45.1	123	1	HV24_MOUSE
66	290.5	44.9	119	1	HV3N_HUMAN
67	288.5	44.6	114	1	HV01_CANFA
68	288	44.5	117	1	HV42_MOUSE
69	288	44.5	140	1	HV02_MOUSE
70	287.5	44.4	119	1	HV3P_HUMAN
71	287	44.4	121	1	HV01_MOUSE
72	287	44.4	144	1	HV26_MOUSE
73	285.5	44.1	111	1	HV35_MOUSE
74	285.5	44.1	116	1	HV1A_RABIT
75	285	44.0	120	1	HV03_MOUSE
76	283	43.7	117	1	HV12_MOUSE
77	281.5	43.5	120	1	HV50_MOUSE
78	281	43.4	113	1	HV34_MOUSE
79	280	43.3	117	1	HV13_MOUSE
80	279.5	43.2	117	1	HV3O_HUMAN
81	279.5	43.2	137	1	HV11_MOUSE
82	278	43.0	136	1	HV15_MOUSE
83	276.5	42.7	117	1	HV55_MOUSE
84	276.5	42.7	118	1	HV3V_HUMAN
85	274	42.3	97	1	HV56_MOUSE
86	271.5	42.0	98	1	HV57_MOUSE
87	270.5	41.8	117	1	HV54_MOUSE
88	269	41.6	116	1	HV3R_HUMAN
89	268.5	41.5	124	1	HV1D_HUMAN
90	266.5	41.2	116	1	HV36_MOUSE
91	265.5	41.0	117	1	HV14_MOUSE
92	264.5	40.9	117	1	HV59_MOUSE
93	264	40.8	125	1	HV1F_HUMAN
94	262	40.5	116	1	HV3Q_HUMAN
95	261.5	40.4	117	1	HV52_MOUSE
96	261.5	40.4	117	1	HV58_MOUSE
97	256.5	39.6	117	1	HV04_MOUSE
98	256.5	39.6	147	1	HVIC_HUMAN
99	255.5	39.5	117	1	HV03_CARAU
100	252.5	39.0	117	1	HV1B_HUMAN

ALIGNMENTS

RESULT 1

```
HV43_MOUSE STANDARD; PRT; 144 AA.
ID HV43_MOUSE
AC P01819;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ig heavy chain V region MOPC 141 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81012133; PubMed=6774258;
RA Sakano H., Maki R., Kurosawa Y., Roeder W., Tonegawa S.;
RT "Two types of somatic recombination are necessary for the generation
of complete immunoglobulin heavy-chain genes.";
RL Nature 286:676-683(1980).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V00757; CAA24148.1; -
DR PIR; A02095; HWS14.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 115 IG HEAVY CHAIN V REGION PJ14.
FT NON_TER 115 115
FT SEQUENCE 115 AA; 12447 MW; 7569DDA4A483D500 CRC64;
SQ
Query Match 67.1%; Score 434; DB 1; Length 115;
Best Local Similarity 85.4%; Pred. No. 6.2e-38;
Matches 82; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHVRPPGKGLWLGIMGGNTDYN 60
Db 20 QVQLKESGPGLVAPQSLSITCTVSGFSLTGTVGNVRPPGKGLWLGIMGGSDTIN 79
QY 61 SALKSRLSTSKNSKNQVFLKNSLTAAATVAVYCA 96
Db 80 SALKSRLSTSKNSKSQVFLKNSLTQDDTARYCA 115
RESULT 3
HV45_MOUSE STANDARD; PRT; 116 AA.
ID HV45_MOUSE
AC P01821;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region MC101 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82075900; PubMed=6273429;
RA Kataoka T., Nikaudo T., Miyata T., Moriwaki K., Honjo T.;
RT "The nucleotide sequences of rearranged and germline immunoglobulin
VH genes of a mouse myeloma MC101 and evolution of VH genes in
mouse.";
RL J. Biol. Chem. 257:277-285(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00502; AAA38515.1; -
DR PIR; A02096; GIMS10.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
```

```
HV44_MOUSE STANDARD; PRT; 115 AA.
ID HV44_MOUSE
AC P01820;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region PJ14 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81012133; PubMed=6774258;
RA Sakano H., Maki R., Kurosawa Y., Roeder W., Tonegawa S.;
RT "Two types of somatic recombination are necessary for the generation
of complete immunoglobulin heavy-chain genes.";
RL Nature 286:676-683(1980).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V00758; CAA24149.1; -
DR PIR; A02094; GZMS14.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 144 IG HEAVY CHAIN V REGION MOPC 141.
FT NON_TER 144 144
FT SEQUENCE 144 AA; 15759 MW; 8E47A7CB3706D30A CRC64;
SQ
Query Match 74.5%; Score 482; DB 1; Length 144;
Best Local Similarity 74.4%; Pred. No. 9.1e-43;
Matches 93; Conservative 9; Mismatches 19; Indels 4; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHVRPPGKGLWLGIMGGNTDYN 60
Db 20 QVQLKESGPGLVAPQSLSITCTVSGFSLTGTVGNVRPPGKGLWLGIMGGSDTIN 79
QY 61 SALKSRLSTSKNSKNQVFLKNSLTAAATVAVYCA 116
Db 80 STLSRLSTSKNSKSQVFLKNSLTQDDTARYCASVSIYYGRSDKYFTLDYWGQGT 139
QY 117 VTSS 121
Db 140 VTSS 144
RESULT 2
HV44_MOUSE STANDARD; PRT; 115 AA.
ID HV44_MOUSE
AC P01820;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region PJ14 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```

```

DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 19
FT CHAIN 20 116 IG HEAVY CHAIN V REGION MC101.
FT NON_TER 116
SO SEQUENCE 116 AA; 12593 MW; 8079A6EE7C552B3E CRC64;

Query Match
Best Local Similarity 65.4%; Score 423; DB 1; Length 116;
Matches 77; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Oy 1 QVQLKESGPGLVKPSQTLSTICTSGFSLRSYVHVRQPPGKGLWGLMGGNTDYN 60
Db 20 QVQLKESGPGLVKPSQTLSTICTSGFSLRSYVHVRQPPGKGLWGLMGGNTDYN 79
Oy 61 SALKSRISKDNKNOVFLKMNLSLTAAADTAVYYCAR 97
80 AAFISRLSISKDNKNOVFLKMNLSLTAAADTAVYYCAR 116

RESULT 4
HV02_XENLA STANDARD; PRT; 135 AA.
ID HV02_XENLA
AC P20957;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region XIG14 precursor (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88176921; PubMed=2451244;
RA Schwager J., Mikoryak C.A., Steiner L.A.;
RT "Amino acid sequence of heavy chain from xenopus laevis Igm deduced
from cDNA sequence: implications for evolution of immunoglobulin
domains".
RL Proc. Natl. Acad. Sci. U.S.A. 85:2245-2249(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03632; AAA49791.1;
DR PIR: B31933; B31933.
DR HSSP: P01810; 2FBJ.
DR InterPro: IPR003006; Ig_MHC.
DR SMART: SM00406; IGV.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT NON_TER 1
FT SIGNAL <1 18
FT CHAIN 19 135 IG HEAVY CHAIN V REGION XIG14.
FT NON_TER 135
SO SEQUENCE 135 AA; 15080 MW; EBC467105C00732E CRC64;

Query Match
Best Local Similarity 61.3%; Score 396.5; DB 1; Length 135;
Matches 72; Conservative 23; Mismatches 20; Indels 3; Gaps 2;

Oy 4 LKESGPGLVKPSQTLSTICTSGFSLRSYVHVRQPPGKGLWGLMGGNTDYN 63
Db 21 LQESGPGTVKPSRLTCTSGFSLRSYVHVRQPPGKGLWGLMGGNTDYN 80

```

```

Oy 64 KSRISKDNKNOVFLKMNLSLTAAADTAVYYCARKEGFYVYDGFVWGQGLTVTS 121
Db 81 KNRVTITKDNCKKQVYLOMNGMEVKDTAMYYCAR--EYASGYN-FDYWGQGTMTVTS 135

RESULT 5
HV3G_HUMAN STANDARD; PRT; 122 AA.
ID HV3G_HUMAN
AC P01768;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region CAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=81013859; PubMed=6774332;
RA Lehman D.W., Putnam F.W.;
RT "Amino acid sequence of the variable region of a human mu chain:
location of a possible JH segment.".
RL Proc. Natl. Acad. Sci. U.S.A. 77:3239-3243(1980).
CC -!- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM THE PLASMA OF A
PATIENT WITH MACROGLOBULINEMIA.
DR PIR: A02051; M3HUAM.
DR HSSP: P01772; 2FB4.
DR InterPro: IPR003006; Ig_MHC.
DR SMART: SM00406; IGV.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT MOD_RES 1
FT NON_TER 122
SO SEQUENCE 122 AA; 13668 MW; A42D0F17D252F1C2 CRC64;

Query Match
Best Local Similarity 57.9%; Score 374.5; DB 1; Length 122;
Matches 68; Conservative 24; Mismatches 29; Indels 1; Gaps 1;

Oy 1 QVQLKESGPGLVKPSQTLSTICTSGFSLRSYVHVRQPPGKGLWGLMGI-WGGGNTDY 59
Db 1 QVELVESGGGVVZPGRSLRLSCAASGFTFSNYSMMHVRQPPGKGLWVAVISYGBBKYY 60
Oy 60 NSALKSRISKDNKNOVFLKMNLSLTAAADTAVYYCARKEGFYVYDGFVWGQGLTVTV 119
Db 61 ABSYKGRFTISRDSKSTLYLQMNLSRAETAVYYCARDRLPLYGBYAFNYWGQGLTVTV 120
Oy 120 SS 121
Db 121 SS 122

RESULT 6
HV2I_HUMAN STANDARD; PRT; 146 AA.
ID HV2I_HUMAN
AC P06331;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-II region ARH-77 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85205332; PubMed=3922855;
RA Kudo A., Ishihara T., Nishimura Y., Watanabe T.;
RT "A cloned human immunoglobulin heavy chain gene with a novel direct-
repeat sequence in 5' flanking region.";

```



```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RA MEDLINE=74175307; PubMed=4208843;
RX Florent G., Lehman D., Putnam F.W.;
RT "The switch point in mu heavy chains of human IgM immunoglobulins.";
RL Biochemistry 13:2482-2498(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A WALDENSTROM'S
CC MACROGLOBULIN.
DR PIR: A02052; M3HUGA.
DR HSSP: P01772; 2F84.
DR DR InterPro; IPR003006; Ig_MHC.
DR DR InterPro; IPR003596; Ig_v.
DR DR Pfam; PF00047; Ig; 1.
DR DR SMART; SM00406; IGV; 1.
DR KW Immunoglobulin V region.
FT MOD_RES 1 1
FT NON_TER 122 122 PYRROLIDONE CARBOXYLIC ACID.
FT SEQUENCE 122 AA; 13166 MW; 74E5B6959E84100A CRC64;
Query Match 54.0%; Score 349.5; DB 1; Length 122;
Best Local Similarity 51.6%; Pred. No. 3.4e-29;
Matches 63; Conservative 27; Mismatches 31; Indels 1; Gaps 1;

OY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHVRQPPGKGLWGLMIW-GGGNTDY 59
DB 1 QVQLVQAGGAVPQGRSLRLSCIASGFTFSNYGMHVRQAPGKGLWVAVIYNGSRYY 60
OY 60 NSALKSRSLTSKDNKQVFLKMNLSLAADTAVYICARKGEFFYGYDGFVYWGOTLVTV 119
DB 61 GDSYKGRFTISRDNKSRKRLTYMZMNSLRTEDTAVYICARDPDILTAF-SFDYWGQGVLT 119
OY 120 SS 121
DB 120 SS 121

RESULT 10
HV2G_HUMAN STANDARD; PRT; 117 AA.
ID HV2G_HUMAN
AC P01825;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-II region NEMM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77243302; PubMed=407927;
RA Poljak R.J., Nakashima Y., Chen B.L., Konigsberg W.;
RT "Amino acid sequence of the VH region of a human myeloma
RT immunoglobulin (IgG New).";
RL Biochemistry 16:3412-3420(1977).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF FAB FRAGMENT.
RX MEDLINE=78066916; PubMed=618887;
RA Saul F.A., Amzel L.M., Poljak R.J.;
RT "Preliminary refinement and structural analysis of the Fab fragment
RT from human immunoglobulin new at 2.0-A resolution.";
RL J. Biol. Chem. 253:585-597(1978).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
DR PIR: A02100; GIHUNM.
DR PDB: 7FAB; 31-JAN-94.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR KW Immunoglobulin V region; 3D-structure.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT STRAND 3 7
FT STRAND 11 12
FT TURN 14 15
FT TURN 18 25
FT TURN 30 31
FT STRAND 33 39
FT TURN 41 42
FT STRAND 46 51
FT TURN 53 54
FT STRAND 57 59
FT TURN 61 63
FT HELIX 64 66
FT TURN 67 72
FT TURN 73 76
FT STRAND 77 82
FT STRAND 87 89
FT STRAND 91 98
FT STRAND 104 107
FT STRAND 111 115
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12790 MW; 2DA47B509562D237 CRC64;
Query Match 53.5%; Score 346; DB 1; Length 117;
Best Local Similarity 56.6%; Pred. No. 7.4e-29;

OY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHVRQPPGKGLWGLMIW-GGGNTDY 59
DB 1 QVQLVQAGGAVPQGRSLRLSCIASGFTFSNYGMHVRQAPGKGLWVAVIYNGSRYY 60
OY 60 NSALKSRSLTSKDNKQVFLKMNLSLAADTAVYICARKGEFFYGYDGFVYWGOTLVTV 119
DB 61 GDSYKGRFTISRDNKSRKRLTYMZMNSLRTEDTAVYICARDPDILTAF-SFDYWGQGVLT 119
OY 120 SS 121
DB 120 SS 121

RESULT 10
HV3J_HUMAN STANDARD; PRT; 121 AA.
ID HV3J_HUMAN
AC P01771;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region HIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=79124695; PubMed=420800;
RA Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;
RT "Amino acid sequence of the VH region of human myeloma
RT cryoimmunoglobulin IgG HIL.";
RL Biochemistry 18:553-560(1979).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
DR PIR: A02054; GIHULH.
DR HSSP: P01772; 2F84.
DR DR InterPro; IPR003006; Ig_MHC.
DR DR InterPro; IPR003596; Ig_v.
DR DR Pfam; PF00047; Ig; 1.
DR DR SMART; SM00406; IGV; 1.
DR KW Immunoglobulin V region.
FT MOD_RES 1 1
FT NON_TER 121 121 PYRROLIDONE CARBOXYLIC ACID.
FT SEQUENCE 121 AA; 13566 MW; 480FC53610EF5DAB CRC64;
Query Match 53.9%; Score 349; DB 1; Length 121;
Best Local Similarity 52.5%; Pred. No. 3.8e-29;
Matches 64; Conservative 23; Mismatches 33; Indels 2; Gaps 2;
```

Matches 69; Conservative 22; Mismatches 25; Indels 6; Gaps 2;

Qy 1 QVQLKESGPGLVKPSQSLTISCTISGFSLSRYSVHVRQPPGKGLWGLMIWGGNTDYN 60
 Db 1 QVQLKESGPGLVKPSQSLTISCTISGFSLSRYSVHVRQPPGKGLWGLMIWGGNTDYN 60
 Qy 61 SALKSLISKDNKNOVFLKMSLTAAADTAVVYCARKEGFFYVYDGFV-YWGQGLTVTV 119
 Db 61 TPLSRVTMLVDTSKNOFSLRSLSSVTAADTAVVYCARN-----LIAGCIDWVGQGLTVTV 115
 Qy 120 SS 121
 Db 116 SS 117

RESULT 11
 HV01_XENLA STANDARD; PRT; 136 AA.
 ID HV01_XENLA
 AC P20956;
 DT 01-FEB-1991 (Rel. 17; Created)
 DT 01-FEB-1991 (Rel. 17; Last sequence update)
 DT 15-JUL-1999 (Rel. 38; Last annotation update)
 DE Ig heavy chain V region XIG8 precursor (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 CC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88176921; PubMed=2451244;
 RA Schwager J., Mikoryak C.A., Steiner L.A.;
 RT "Amino acid sequence of heavy chain from Xenopus laevis IgM deduced
 from cDNA sequence: implications for evolution of immunoglobulin
 RT domains";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:2245-2249(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M20484; AAA49774.1; ALT_TERM.
 DR PIR; A31933; A31933.
 DR HSSP; P01825; 7FAB.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR Immunoglobulin V region; Signal.
 DR NON_TER 1
 DR SIGNAL <1 18
 DR FT CHAIN 19 136 IG HEAVY CHAIN V REGION XIG8.
 DR FT NON_TER 136 136
 DR SQ SEQUENCE 136 AA; 15123 MW; 3141838981441963 CRC64;
 Query Match 53.2%; Score 344; DB 1; Length 136;
 Best Local Similarity 53.4%; Pred. No. 1.4e-28;
 Matches 63; Conservative 24; Mismatches 29; Indels 2; Gaps 1;

Qy 4 LKESGPGLVKPSQSLTISCTISGFSLSRYSVHVRQPPGKGLWGLMIWGGNTDYNAL 63
 Db 21 LQESGPGTVKPSQSLTISCTISGFSLSRYSVHVRQPPGKGLWGLMIWGGNTDYNAL 80
 Qy 64 KSLRSLKDNKNOVFLKMSLTAAADTAVVYCARKEGFFYVYDGFV-YWGQGLTVTVS 121
 Db 81 KNRVTITKDNKQVQLQMGMEVKDAMYCYCTSTAGTAGY--FEHWGQGTMTVTVS 136

RESULT 12.

HV2C_HUMAN STANDARD; PRT; 119 AA.
 ID HV2C_HUMAN
 AC P01816;
 DT 21-JUL-1986 (Rel. 01; Created)
 DT 21-JUL-1986 (Rel. 01; Last sequence update)
 DT 15-JUL-1999 (Rel. 38; Last annotation update)
 DE Ig heavy chain V-II region DAW.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=70258837; PubMed=5449120;
 RA Press E.M., Hogg N.M.;
 RT "The amino acid sequences of the Fd fragments of two human gamma-1
 RT heavy chains";
 RL Biochem. J. 117:641-660(1970).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM IGG1 ISOLATED FROM THE
 CC SERUM OF A PATIENT WITH HYPERGAMMAGLOBULINEMIA.
 DR PIR; A02091; G1HUDW.
 DR HSSP; P01789; 1MCP.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR Immunoglobulin V region.
 DR MOD_RES 1
 DR NON_TER 119 119 PYRROLIDONE CARBOXYLIC ACID.
 DR SQ SEQUENCE 119 AA; 13045 MW; 4E13E00214BAD789 CRC64;
 Query Match 52.9%; Score 342; DB 1; Length 119;
 Best Local Similarity 57.6%; Pred. No. 1.9e-28;
 Matches 72; Conservative 15; Mismatches 28; Indels 10; Gaps 4;

Qy 1 QVQLKESGPGLVKPSQSLTISCTISGFSLS--RYSVHVRQPPGKGLWGLMIWGGNTD 58
 Db 1 QVTLRESGPALVREPTQTLLTCTFSGLSGETMCVAVIROPPEALEWLA--WDILNDD 58
 Qy 59 --YNSALKSRSLKDNKNOVFLKMSLTAAADTAVVYCARKEGFFYVYDGFVYWGQGL 116
 Db 59 KYCGASLETRVLAVSKDYSKNOVLSMTVPGDTATYYCARS---CGSQYFDYWGQGL 114
 Qy 117 VTSS 121
 Db 115 VTSS 119

RESULT 13
 HV61_MOUSE STANDARD; PRT; 116 AA.
 ID HV61_MOUSE
 AC P18532;
 DT 01-NOV-1990 (Rel. 16; Created)
 DT 01-NOV-1990 (Rel. 16; Last sequence update)
 DT 15-JUL-1999 (Rel. 38; Last annotation update)
 DE Ig heavy chain V region 1B43 precursor.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/CJ;
 RX MEDLINE=89279149; PubMed=2499654;
 RA Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;
 RT "Early onset of somatic mutation in immunoglobulin VH genes during
 RT the primary immune response";
 RL J. Exp. Med. 169:2007-2019(1989).
 CC -I- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH3660 SUBFAMILY.
 DR PIR; J0508; HVMS1B.
 DR HSSP; P01825; 7FAB.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.

Db 121 OCTTVHVSS 129

RESULT 15

HV3K_HUMAN

ID HV3K_HUMAN STANDARD; PRT; 126 AA.

AC P01772;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig heavy chain V-III region KOL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

[1]

RP SEQUENCE, AND DISULFIDE BONDS.

RX MEDLINE=83289131; PubMed=6884994;

RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;

RT "Three-dimensional structure determination of antibodies. Primary

RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.,"

RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).

RX MEDLINE=81072295; PubMed=7441755;

RA Marquart M., Deisenhofer J., Huber R., Palm W.;

RT "Crystallographic refinement and atomic models of the intact

RT immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A

and 1.0-A resolution.,"

FT J. Mol. Biol. 141:369-391(1980).

RL PIR; A02055; GIHUKL.

DR PDB; 2FB4; 12-JUL-89.

DR PDB; 2IG2; 12-JUL-89.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003596; Ig_V.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; IGV; 1.

DR Immunoglobulin V region; 3D-structure.

KW MOD_RES 1 1

FT PYRROLIDONE CARBOXYLIC ACID.

FT DISULFID 22 96

FT DISULFID 105 110

FT STRAND 3 7

FT STRAND 11 12

FT TURN 14 15

FT STRAND 18 25

FT STRAND 29 31

FT HELIX 34 39

FT STRAND 41 42

FT TURN 46 51

FT STRAND 53 54

FT TURN 58 60

FT STRAND 62 67

FT TURN 68 73

FT TURN 74 77

FT STRAND 78 83

FT HELIX 88 90

FT STRAND 92 99

FT STRAND 106 106

FT TURN 107 108

FT STRAND 109 109

FT STRAND 113 113

FT STRAND 120 124

FT STRAND 126 126

FT NON_TER 126

SQ SEQUENCE 126 AA; 13718 MW; E4D71B52B16F8776 CRC64;

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYVROAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similar

Db 121 OCTTVHVSS 129

RESULT 15

HV3K_HUMAN

ID HV3K_HUMAN STANDARD; PRT; 126 AA.

AC P01772;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig heavy chain V-III region KOL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

1

SEQUENCE, AND DISULFIDE BONDS.

RX MEDLINE=83289131; PubMed=6884994;

RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;

RT "Three-dimensional structure determination of antibodies. Primary

RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.,"

RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).

RX MEDLINE=81072295; PubMed=7441755;

RA Marquart M., Deisenhofer J., Huber R., Palm W.;

RT "Crystallographic refinement and atomic models of the intact

RT immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A

and 1.0-A resolution.,"

FT J. Mol. Biol. 141:369-391(1980).

RL PIR; A02055; GIHUKL.

DR PDB; 2FB4; 12-JUL-89.

DR PDB; 2IG2; 12-JUL-89.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003596; Ig_V.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; IGV; 1.

DR Immunoglobulin V region; 3D-structure.

KW MOD_RES 1 1

FT PYRROLIDONE CARBOXYLIC ACID.

FT DISULFID 22 96

FT DISULFID 105 110

FT STRAND 3 7

FT STRAND 11 12

FT TURN 14 15

FT STRAND 18 25

FT STRAND 29 31

FT HELIX 34 39

FT STRAND 41 42

FT TURN 46 51

FT STRAND 53 54

FT TURN 58 60

FT STRAND 62 67

FT TURN 68 73

FT TURN 74 77

FT STRAND 78 83

FT HELIX 88 90

FT STRAND 92 99

FT STRAND 106 106

FT TURN 107 108

FT STRAND 109 109

FT STRAND 113 113

FT STRAND 120 124

FT STRAND 126 126

FT NON_TER 126

SQ SEQUENCE 126 AA; 13718 MW; E4D71B52B16F8776 CRC64;

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity 51.9%; Pred. No. 9.7e-28;

Matches 68; Conservative 22; Mismatches 26; Indels 15; Gaps 4

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRQPPGKGLWLGNIW-GGNTDY 59

Db 1 QVQLVESGGGVQPGSRSLRSCSSSGIFSSYAMVYWRQAPGKGLWLVINDDGSDQHY 60

Query Match 51.9%; Score 335.5; DB 1; Length 126;

Best Local Similarity

```

QY 60 NSALKSRISKDNKNOVFLKMNLSLTADTAVYYCARKGEFYVDGFFV-----Y 110
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 ADVKRGRTISRDNKNTLFLQMSLRPEDTGVYFCARDG-----GH-GFCSSASCFGPDY 115

QY 111 WQGGTLVTVSS 121
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 WQGGTPVTVSS 126

RESULT 16
HV47_MOUSE
ID HV47_MOUSE STANDARD; PRT; 113 AA.
AC P01823;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 36-60.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE.
RC STRAIN=A/J;
RX MEDLINE=64024551; PubMed=6414509;
RA Juszczak E.C., Margolies M.N.;
RT "Amino acid sequence of the heavy chain variable region from the A/J
RT mouse anti-arsenate monoclonal antibody 36-60 bearing a minor
RT idiotypic.";
RL Biochemistry 22:4291-4296(1983).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN ANTIARSONATE
CC MONOCLONAL ANTIBODY OF THE IGG2A SUBCLASS. IT REPRESENTS A SECOND
CC IDIOTYPE FAMILY CHARACTERISTIC OF THE ANTIARSONATE RESPONSE OF
CC STRAIN A/J MICE.
DR PIR; A02098; G2MS60.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Antiarsenate antibody.
FT NON_TER 113
FT SEQUENCE 113 AA; 12734 MW; 38DC0E0E3F5075B7 CRC64;

Query Match 51.5%; Score 333; DB 1; Length 113;
Best Local Similarity 53.7%; Pred. No. 1.5e-27;
Matches 65; Conservative 23; Mismatches 25; Indels 8; Gaps 1;

1 QVQLKESGPGLVKPSQTLISCTISGFSLSRYSVHVRPQPGKGLWLGMIWGGNTDYN 60
1 EVQLQESGPGSLVKPSQTLISCTISGFSLSRYSVHVRPQPGKGLWLGMIWGGNTDYN 60

61 SALKSRISKDNKNOVFLKMNLSLTADTAVYYCARKGEFYVDGFFVWQGGTLVTVS 120
61 PSLSRSIRITRDTSKNQYQLQNSVTSEDATYYCTSL-----RFAYWQGGTLVTVS 112

121 S 121
113 A 113

RESULT 17
HV31_HUMAN
ID HV31_HUMAN STANDARD; PRT; 119 AA.
AC P01770;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region NIE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE.
RX MEDLINE=77070269; PubMed=826475;
RA Ponstingl H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a
RT monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The
RT chymotryptic peptides of the H-chain, alignment of the tryptic
RT peptides and discussion of the complete structure.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [2]
RP DISULFIDE BOND.
RX MEDLINE=77070267; PubMed=1002129;
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT "Rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie). I: Purification and
RT characterization of the protein, the L- and H-chains, the
RT cyanogen bromide cleavage products, and the disulfide bridges.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
DR PIR; A02053; GHUNI.
DR HSSP; P01772; 2FBA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region.
FT MOD_RES 1
FT DISULFID 22 96
FT NON_TER 119
FT SEQUENCE 119 AA; 13242 MW; C96935A6E55E165B CRC64;

Query Match 51.5%; Score 333; DB 1; Length 119;
Best Local Similarity 52.0%; Pred. No. 1.6e-27;
Matches 65; Conservative 22; Mismatches 28; Indels 10; Gaps 3;

QY 1 QVQLKESGPGLVKPSQTLISCTISGFSLSRYSVHVRPQPGKGLWLG-MIWGGNTDY 59
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 QVQLVQSGGPGVQPGKGLWLGSCAASGFTFSRTIHWVQAPGKGLWVAVMSYBGBKH 50

QY 60 NSALKSRISKDNKNOVFLKMNLSLTADTAVYYCAR---KGEFYVDGFFVWQGGTL 116
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 ADVKRGRTISRDNKNTLFLQMSLRPEDTAVYYCARIDTAME-----FAHWGQGT 114

QY 117 VTVSS 121
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 VTVSS 119

RESULT 18
HV40_MOUSE
ID HV40_MOUSE STANDARD; PRT; 119 AA.
AC P01810;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region J539.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
PRELIMINARY SEQUENCE.
RX MEDLINE=79223895; PubMed=111245;
RA Rao D.N., Rudikoff S., Krutzsch H., Potter M.;
RT "Structural evidence for independent joining region gene in
RT immunoglobulin heavy chains from anti-galactan myeloma proteins and
RT its potential role in generating diversity in
RT complementarity-determining regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894(1979).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=88217852; PubMed=3449853;

```

```

RA Suh S.W., Bhat T.N., Navia M.A., Cohen G.H., Rao D.N., Rudikoff S.,
RA Davies D.R.;
RT "The galactan-binding immunoglobulin Fab J539: an X-ray diffraction
RL study at 2.6-A resolution.";
CC Proteins 1:74-80(1986).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS GALACTAN.
CC PIR: A02080; AVMSJ5.
DR PDB: 2FBJ; 15-OCT-90.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; 3D-structure.
FT NON_TER 119
FT STRAND 3
FT TURN 7
FT STRAND 10
FT TURN 12
FT STRAND 14
FT TURN 15
FT STRAND 18
FT TURN 25
FT STRAND 29
FT TURN 31
FT STRAND 34
FT TURN 39
FT STRAND 41
FT TURN 42
FT STRAND 45
FT TURN 51
FT STRAND 53
FT TURN 54
FT STRAND 58
FT TURN 60
FT STRAND 62
FT TURN 67
FT STRAND 68
FT TURN 72
FT STRAND 78
FT TURN 83
FT STRAND 88
FT TURN 90
FT STRAND 92
FT TURN 100
FT STRAND 101
FT TURN 103
FT STRAND 104
FT TURN 108
FT STRAND 112
FT TURN 116
SQ SEQUENCE 119 AA; 13240 MW; 577B4F1DB675C1F1 CRC64;

Query Match 51.48; Score 332.5; DB 1; Length 119;
Best Local Similarity 52.58; Pred. No. 1.8e-27;
Matches 64; Conservative 23; Mismatches 30; Indels 5; Gaps 3;

Qy 1 QVQLKESGPGLVKPSQTLTCTISGFSLSRYSVHVRQPPGKGLVGLMIW-CGGNTDY 59
Db 1 EVKLESGGGLVPGGSLKLSCAASGDFSKYWNWSVVRQAPKGLWGEIHPDSGTINY 60

Qy 60 NSALKSRLSISKNSKNOVFLKMNLSLTADTAVYICARKGEFYGYDGFVYWGQTLTV 119
Db 61 TPSLKDRIISRDNAKNSLYLQMSKVRSEDTALYICARL--HYGYN--AYWGQTLTV 116

120 SS 121
117 SA 118

RESULT 19
HV16_MOUSE
ID HV16_MOUSE STANDARD; PRT; 136 AA.
AC P01783;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region MOPC 21 precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81234548; PubMed=6789376;
RA Bochwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPb family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
RN [2]

SEQUENCE OF 17-136.
MEDLINE=77100368; PubMed=401950;
RA Adelugbo K., Milstein C., Secher D.S.;
RT "Molecular analysis of spontaneous somatic mutants.";
RL Nature 265:299-304(1977).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to a license@sib-sib.ch).
CC -----
CC EMBL: J00522; AAD15290.1; -
CC PIR: A02066; GLMS21.
CC HSSP: P01772; 2FB4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT NON_TER 1
FT SIGNAL <1 16
FT CHAIN 17 136 IG HEAVY CHAIN V REGION MOPC 21.
FT DOMAIN 115 119 D SEGMENT.
FT DOMAIN 120 136 JH4 SEGMENT.
FT DISULFID 38 112
FT CONFLICT 75 78
FT CONFLICT 89 90
FT CONFLICT 115 115 DN -> ND (IN REF. 2).
FT CONFLICT 120 120 W -> H (IN REF. 2).
FT NON_TER 136 Y -> W (IN REF. 2).
SQ SEQUENCE 136 AA; 15071 MW; 2276A98DBDBF7016 CRC64;

Query Match 50.98; Score 329.5; DB 1; Length 136;
Best Local Similarity 54.18; Pred. No. 4.4e-27;
Matches 66; Conservative 17; Mismatches 34; Indels 5; Gaps 3;

Qy 2 VOLKESGPGLVKPSQTLTCTISGFSLSRYSVHVRQPPGKGLVGLMIWGGNT-DYN 60
Db 18 VOLVESGGGLVPGGSKLSCAASGFTFSFGMHVVRQAPKGLWYVSSGSTLIYA 77

Qy 61 SALKSRLSISKNSKNOVFLKMNLSLTADTAVYICARKGEF-YGYDGFVYWGQTLTV 119
Db 78 DTVKGRFTISRDNPKNLTFLQMTSLRSEDTAMYICARWNPYPYAMD--VWGQTSVT 134

120 SS 121
135 SS 136

RESULT 20
HV37_MOUSE
ID HV37_MOUSE STANDARD; PRT; 119 AA.
AC P01807;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region X44.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79223895; PubMed=111245;
RA Rao D.N., Rudikoff S., Krutzsch H., Potter M.;
RT "Structural evidence for independent joining region gene in
RT immunoglobulin heavy chains from anti-galactan myeloma proteins and
RT its potential role in generating diversity in
RT complementarity-determining regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894(1979).

```



```

Db      1  EVKLLSEGGGLVQPGGSL-USCAASGPDFSRYWMSWARPAGKQEWIGINPGSSITINY 59
QY      60  NSALKSRLSISKDNSKNOVFLKMNSLTAADTAVYCYARKGEFYGYDGFFVWGGQTLVTV 119
Db      60  TPSLKDKEFIISRDNAKNTLYLQMSKVRSEDTALYYCARLG--YVGY--FDYWGOGTTLTV 115
QY      120  SS 121
Db      116  SS 117

RESULT 25
HV01_RAT
ID HV01_RAT STANDARD; PRT; 142 AA.
AC P01805;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain v region IR2 precursor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83064537; PubMed=6292865;
RA Hellman L., Pettersson U., Engstrom A., Karlsson T., Bennich H.;
RT "Structure and evolution of the heavy chain from rat immunoglobulin
RT E.";
RL Nucleic Acids Res. 10:6041-6049(1982).
CC -! IMMUNOCYTOMA: THE MRNA WAS ISOLATED FROM AN IGE-SECRETING
CC IMMUNOCYTOMA THAT ARISES SPONTANEOUSLY IN LOU/C/WSL RATS.
DR PIR; A02075; EVRTR2.
DR HSP; P01789; 1MCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region; Signal.
KW SIGNAL
FT CHAIN 1
FT CHAIN 19
FT NON_TER 20 142 IG HEAVY CHAIN V REGION IR2.
FT NON_TER 142 142
SQ SEQUENCE 142 AA; 15024 MW; DE29E6CPE745DF3B CRC64;

Query Match 49.8%; Score 322; DB 1; Length 142;
Best Local Similarity 48.8%; Pred. No. 2.7e-26;
Matches 63; Conservative 24; Mismatches 28; Indels 14; Gaps 3;

QY      1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWYRQPPGKGLGMIWGNGN---T 57
Db      19  EVKLLSEGGGLVQPGGMSVKLSCATSGFTSDYNMEWYRQAPGKGLWVAEIRKANNYYA 79
QY      58  DYNALKSRLSISKDNSKNOVFLKMNSLTAADTAVYCYARKGEFYGYDG-----FVYWG 112
Db      80  YYGKSLKGRFTLSRDSKSIYVLQMNIRSDGTGIYYCSR-----GYGGYSENWFWYWG 133
QY      113  QGTLVTVSS 121
Db      134  QGTLVTVSS 142

Search completed: June 23, 2003, 14:02:09
Job time : 4.50532 secs

```

Search completed: June 23, 2003, 14:02:09
Job time : 4.50532 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:10 ; Search time 15.6613 Seconds
(without alignments)
1591.926 Million cell updates/sec

Title: US-10-056-052a-20

Perfect score: 647

Sequence: 1 QVOKSGPGLVKPQSOTLSI.....YGYDGFVWVGOTLTVTSS 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

671580 seqs, 206047115 residues

number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	465.5	71.9	121	11 Q99NG4	Q99ng4 mus musculus
2	463.5	71.6	482	11 Q91X92	Q91x92 mus musculus
3	413	63.8	119	4 Q9UL73	Q9ul73 homo sapien
4	402.5	62.2	613	4 Q96EY0	Q96ey0 homo sapien
5	385.5	59.6	496	4 Q96KX8	Q96kx8 homo sapien
6	373.5	57.7	150	4 Q95973	Q95973 homo sapien
7	369	57.0	479	11 Q99M22	Q99m22 mus musculus
8	363.5	56.2	573	4 Q8WU38	Q8wu38 homo sapien
9	363	56.1	121	4 Q9UL96	Q9ul96 homo sapien
10	359	55.5	121	4 Q9UL71	Q9ul71 homo sapien
11	355.5	54.9	613	4 Q8WUK1	Q8wuk1 homo sapien
12	354	54.7	588	4 Q8WUX4	Q8wux4 homo sapien
13	354	54.7	597	4 Q9BU10	Q9bu10 homo sapien
14	354	54.7	618	4 Q96AA6	Q96aa6 homo sapien
15	351	54.3	116	4 Q9UL93	Q9ul93 homo sapien
16	350.5	54.2	122	4 Q9UL75	Q9ul75 homo sapien

Q9y509 homo sapien
Q9bq88 homo sapien
Q96bb9 homo sapien
Q8r3v9 mus musculus
Q8tc63 homo sapien
Q9ul90 homo sapien
Q9hcc1 homo sapien
Q9ul72 homo sapien
Q920e7 mus musculus
Q9d814 mus musculus
Q9ul88 homo sapien
Q9ul84 homo sapien
Q99ka4 mus musculus
Q96k68 homo sapien
Q8tc77 homo sapien
Q9yvf0 mus musculus
Q9ul91 homo sapien
Q8vea0 mus musculus
Q9n0w6 oryctolagus
Q9n0w4 oryctolagus
Q91z07 mus musculus
Q91lc4 mus musculus
Q99125 mus musculus
Q91205 mus musculus
Q91wt1 mus musculus
Q43234 homo sapien
Q924q7 mus musculus
Q924q1 mus musculus
Q9ul92 homo sapien
Q91wr1 mus musculus
Q9ul87 homo sapien
Q924p7 mus musculus
Q925s3 mus musculus
Q91wp5 mus musculus
Q924r1 mus musculus
Q924q6 mus musculus
Q924r8 mus musculus
Q924q5 mus musculus
Q924r4 mus musculus
Q8vcx7 mus musculus
Q8vj11 mus musculus
Q8vcx4 mus musculus
Q8vcv5 mus musculus
Q9qxf0 mus musculus
Q924q3 mus musculus
Q924r0 mus musculus
Q924q9 mus musculus
Q96qso homo sapien
Q9qxe9 mus musculus
Q924q0 mus musculus
Q920e8 mus musculus
Q920e8 mus musculus
Q924r5 mus musculus
Q91a4 mus musculus
Q924p5 mus musculus
Q924r2 mus musculus
Q9ul74 homo sapien
Q924p8 mus musculus
Q8vdc9 mus musculus
Q8r3h6 mus musculus
Q991a6 mus musculus
Q924r7 mus musculus
Q91va2 mus musculus
Q921k1 mus musculus
Q99131 mus musculus
Q91xe1 mus musculus
Q8wy24 homo sapien
Q924r6 mus musculus
Q9gyz2 schistosoma
Q924q4 mus musculus
Q924p6 mus musculus
Q924r3 mus musculus
Q921c6 mus musculus
Q924q2 mus musculus

90 284.5 44.0 109 11 Q9JL75 Q9j175 mus musculus
91 283.5 43.8 241 11 Q92IA6 Q92IA6 mus musculus
92 281 43.4 125 4 Q9UL95 Q9UL95 homo sapien
93 281 43.4 170 11 Q92S82 Q92S82 mus musculus
94 277.5 42.9 118 11 Q92IC4 Q92IC4 mus musculus
95 277 42.8 614 4 Q96GA6 Q96GA6 homo sapien
96 275 42.5 146 11 Q92Q08 Q92Q08 mus musculus
97 274.5 42.4 143 11 Q924P9 Q924P9 mus musculus
98 274 42.3 119 4 Q9UL94 Q9UL94 homo sapien
99 273.5 42.3 143 11 Q9IV67 Q9IV67 mus musculus
100 273 42.2 112 4 Q9UGP3 Q9UGP3 homo sapien

ALIGNMENTS

```
RESULT 1
Q99NG4 ID Q99NG4 PRELIMINARY; PRT; 121 AA.
Q99NG4 01-JUN-2001 (TReMBLrel. 17, Created)
01-JUN-2001 (TReMBLrel. 17, Last sequence update)
01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Single chain Fv (Fragment).
OS Mus musculus (Mouse).
OG plasmid pHEN1.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=98169018; PubMed=9510199;
RA Hawlich W., Frank R., Hennecke M., Baensch M., Sohns B., Arseniev L.,
RT "Site-Directed C3a-Receptor Antibodies from Phage Display Libraries.";
RL J. Immunol. 160:2947-2958(1998).
DR EMBL; AJ222590; CAA10890.1; -.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR SMART; SM00410; IG_like; 1.
KW Plasmid.
NON_TER 1 1
NON_TER 121 121
SEQUENCE 121 AA; 13255 MW; D293B4EBC8C59D5B CRC64;

Query Match 71.9%; Score 465.5; DB 11; Length 121;
Best Local Similarity 74.0%; Pred. No. 4.9e-44;
Matches 91; Conservative 9; Mismatches 18; Indels 5; Gaps 2;

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLVGLWIGMGNTDYN 60
Db 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLVGLWIGMGNTKYH 60

Qy 61 SALKSRISIKDKNQVFLKMNLSLTAAADTAVYYCARKGEFFYYGDFGV--YWGQGLTVT 118
Db 1 SALKSRISIKDKNQVFLKMNLSLTAAADTAVYYCARKGEFFYYGDFGV--YWGQGLTVT 118

Qy 119 VSS 121
Db 118 VSS 120

RESULT 2
Q91X92 ID Q91X92 PRELIMINARY; PRT; 482 AA.
AC Q91X92.1
```

```
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:18822).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011181; AAH11181.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ SEQUENCE 482 AA; 51865 MW; 312E01F9C1BC7F3C CRC64;

Query Match 71.6%; Score 463.5; DB 11; Length 482;
Best Local Similarity 74.6%; Pred. No. 4.6e-43;
Matches 91; Conservative 11; Mismatches 15; Indels 5; Gaps 2;

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLVGLWIGMGNTDYN 60
Db 20 QVQLKESGPDVAPPSQSLSTICTVSGFALTSAISVVRQPPGKGLVGLVITGVTN 79

Qy 61 SALKSRISIKDKNQVFLKMNLSLTAAADTAVYYCARKGEFFYYGDFGV--YWGQGLTVT 119
Db 80 SALKSRISIKDKNQVFLKMNLSLTNDTARYYCARS----NYEGAMDYWGQGSVT 135

Qy 120 SS 121
Db 136 SS 137

RESULT 3
Q9UL73 ID Q9UL73 PRELIMINARY; PRT; 119 AA.
AC Q9UL73;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Werwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035041; A056277.1; -.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGv; 1.
DR NON_TER 1 1
DR NON_TER 119 119
SQ SEQUENCE 119 AA; 13219 MW; 1BDB86B6420EA0BE CRC64;

Query Match 63.8%; Score 413; DB 4; Length 119;
Best Local Similarity 65.9%; Pred. No. 3.3e-38;
Matches 81; Conservative 16; Mismatches 20; Indels 6; Gaps 2;

Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLVGLWIGMGNTDYN 60
Db 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLVGLWIGMGNTDYN 60
```


Q99M22;
 01-JUN-2001 (TrEMBLrel. 17, Created)
 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 52.0 kDa protein.
 OS Mus musculus (Mouse).
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10990;
 RN [1]
 SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC002091; AA02091.1;
 DR HSSP; P01810; 2FBT
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003597; Ig.cl.
 DR InterPro; IPR003600; Ig_likelike.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 Pfam; PF00047; Ig; 4.
 SMART; SM00409; IG; 3.
 SMART; SM00407; IGcl; 3.
 DR SMART; SM00406; IGV; 1.
 DR SMART; SM00410; IG_likelike; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
 KW Hypothetical protein.
 SQ SEQUENCE 479 AA; 51992 MW; 768E39A138918892 CRC64;
 Query Match 57.0%; Score 369; DB 11; Length 479;
 Best Local Similarity 58.7%; Pred. No. 1.5e-32;
 Matches 71; Conservative 22; Mismatches 22; Indels 6; Gaps 2;
 QY 2 VOLKESGGLVKPSQTLITCTISGFSLSRYSVHWVWVPPGKGLWLGMIWGGNTDYN 60
 DB 20 VOLKESGGLVKPSQTLITCTISGFSLSRYSVHWVWVPPGKGLWLGMIWGGNTDYN 79
 QY 61 SALKSRSLTSKDNKQVFLKMNLSLTAAADTAVYVCARKGEFYVDGFFVYWGQGLTVTS 120
 DB 80 PSLKSRISITRDTSKNQFLKNSVTTEDTATYYCASR-----GYSWFPNGQGLTVTS 134
 QY 121 S 121
 DB 135 A 135
 RESULT 8
 Q8WU38 PRELIMINARY; PRT; 573 AA.
 Q8WU38;
 01-MAR-2002 (TrEMBLrel. 20, Created)
 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 63.0 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 SEQUENCE FROM N.A.
 RC TISSUE=TONSIL;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC021276; AA021276.1;
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003597; Ig.cl.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 Pfam; PF00047; Ig; 4.
 SMART; SM00409; IG; 1.
 SMART; SM00407; IGcl; 3.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.

KW Hypothetical protein.
 SQ SEQUENCE 573 AA; 62967 MW; FD072344033AC530 CRC64;
 Query Match 56.2%; Score 363.5; DB 4; Length 573;
 Best Local Similarity 53.5%; Pred. No. 7.6e-32;
 Matches 68; Conservative 23; Mismatches 27; Indels 9; Gaps 3;
 QY 1 QVQLKESGPGLVKPSQTLITCTISGFSLSRYSVHWVWVPPGKGLWLGMIWGGNTDY 59
 DB 20 EVQLVESGGGLVQPGSRSLRCAASGFTPDYAHMWVWVPPGKGLWVSGISWNSGSGY 79
 QY 60 NSALKSRSLTSKDNKQVFLKMNLSLTAAADTAVYVCARKGE-----FYGYDGEFYWGQ 114
 DB 80 ADSVKGRTISRDNAKNSLYLQMNLSRAEDTALYYCAKHGSGSYGYYYGMD---VMGOG 136
 QY 115 TLTVSS 121
 DB 137 TTVTSS 143
 RESULT 9
 Q9UL96 PRELIMINARY; PRT; 121 AA.
 Q9UL96;
 01-MAY-2000 (TrEMBLrel. 13, Created)
 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Myosin-reactive immunoglobulin heavy chain variable region
 DE (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=98277139; PubMed=9614934;
 RX Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035018; AAD56254.1;
 DR HSSP; P01825; 7FAB.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 FT NON_TER 1
 FT NON_TER 121
 SQ SEQUENCE 121 AA; 13695 MW; D582D450596BDD35 CRC64;
 Query Match 56.1%; Score 363; DB 4; Length 121;
 Best Local Similarity 60.0%; Pred. No. 1.2e-32;
 Matches 75; Conservative 15; Mismatches 27; Indels 8; Gaps 3;
 QY 1 QVQLKESGPGLVKPSQTLITCTISGFSLSRYSVHWVWVPPGKGLWLGMIWGGNTD 58
 DB 1 QITLKESGPTLVKPTQTLTCTISGFSLSLTSGMDVGVIRPPGKALEWLIYDDDKR 60
 QY 59 YNSALKSRSLTSKDNKQVFLKMNLSLTAAADTAVYCA--RKGEFYGYDGEFYWGQGL 116
 DB 61 YPSLKSRLLTITKDTSKNQVLTMTFMDPMDTATYYCAHRSKSGDGY----FDYWGQGL 116
 QY 117 VTSS 121
 DB 117 VTSS 121
 RESULT 10
 Q9UL71 PRELIMINARY; PRT; 121 AA.
 Q9UL71;
 01-MAY-2000 (TrEMBLrel. 13, Created)

```
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035043; RAD56279.1;
DR HSP; P01772; 2F84.
InterPro: IPR003006; Ig_MHC.
Pfam: PF00047; Ig_1.
SMART: SM00406; IgV_1.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13154 MW; 2F045CCFA5D50736 CRC64;

Query Match 55.5%; Score 359; DB 4; Length 121;
Best Local Similarity 57.4%; Pred. No. 3.5e-32;
Matches 70; Conservative 21; Mismatches 29; Indels 2; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSITCTISGFSLSRYSVHVRPPGKGLWLGMIWG-NGNTDY 59
Db :||| ||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
1 EVQLVESGGGVVQPGGSLRLCAASGFTFDGVAHMHVRQAPGKGLWVSLISGDSGNTY 60

QY 60 NSALKSRLSISKDNSKNOVFLKMNLSLTAAATVAVYVCARKGEFYGYDGFVYWGQGLTVTV 119
Db :||| ||| ||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
61 ADSVAGRTISRDNSKNTLYLQMSLRADETAVYYCA-KGKVTITYIDRFIDWGQGTMTVTV 119

QY 120 SS 121
Db ||
120 SS 121

RESULT 11
Q8WUK1 PRELIMINARY; PRT; 613 AA.
ID Q8WUK1;
01-MAR-2002 (TREMBLrel. 20, Created)
01-MAR-2002 (TREMBLrel. 20, Last sequence update)
01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 67.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020240; AAH20240.1;
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
Pfam: PF00047; Ig_5.
SMART: SM00409; Ig; 2.
SMART: SM00407; IgC1; 4.
SMART: SM00406; IgV; 1.
PROSITE: PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671E315 CRC64;
```

```
Query Match 54.9%; Score 355.5; DB 4; Length 613;
Best Local Similarity 55.7%; Pred. No. 6.4e-31;
Matches 68; Conservative 23; Mismatches 28; Indels 3; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSITCTISGFSLSRYSVHVRPPGKGLWLGMIWG-NGNTDY 59
Db :||| ||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
20 QVQLVESGGGVVQPGGSLRLCAASGFTFDGVAHMHVRQAPGKGLWVAVISYDGSNKYY 79

QY 60 NSALKSRLSISKDNSKNOVFLKMNLSLTAAATVAVYVCARKGEFYGYDGFVYWGQGLTVTV 119
Db :||| ||| ||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
80 ADSVAGRTISRDNSKNTLYLQMSLRADETAVYYCAK--DWSEGEVEIFDIWGQGTMTVTV 137

QY 120 SS 121
Db ||
138 SS 139

RESULT 12
Q8WUX4 PRELIMINARY; PRT; 588 AA.
ID Q8WUX4;
01-MAR-2002 (TREMBLrel. 20, Created)
01-MAR-2002 (TREMBLrel. 20, Last sequence update)
01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 64.4 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019235; AAH19235.1;
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
Pfam: PF00047; Ig; 5.
SMART: SM00409; Ig; 2.
SMART: SM00407; IgC1; 4.
SMART: SM00406; IgV; 1.
PROSITE: PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 588 AA; 64438 MW; FC60DBAD82B39FD7 CRC64;

Query Match 54.7%; Score 354; DB 4; Length 588;
Best Local Similarity 57.0%; Pred. No. 8.9e-31;
Matches 73; Conservative 18; Mismatches 27; Indels 10; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSITCTISGFSLSRYSVHVRPPGKGLWLGMIWG-NGNTDY 60
Db :||| ||| ||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
20 QVQLVESGGGVVQPGGSLRLCAASGFTFDGVAHMHVRQAPGKGLWVAVISYDGSNKYY 79

QY 61 SALKSRLSISKDNSKNOVFLKMNLSLTAAATVAVYVCARKGEFYGYDGFVYWGQGLTVTV 113
Db :||| ||| ||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
80 PSLKSRVTISVDTSKKQLSLKSSVNAADTAVYYCARVITRASPGTDGRYGM---VWQ 136

QY 114 GTLTVTVSS 121
Db || |||||
137 GTTTVTVSS 144

RESULT 13
Q9BU10 PRELIMINARY; PRT; 597 AA.
ID Q9BU10;
01-JUN-2001 (TREMBLrel. 17, Created)
01-JUN-2001 (TREMBLrel. 17, Last sequence update)
01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 65.3 kDa protein.
OS Homo sapiens (Human).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002963; AAH02963.1;
DR HSP; P01825; 7EAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; Igc1; 4.
DR SMART; SM00406; Igv; 1.
DR SMART; SM00410; Ig-like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
Hypothetical protein.
SEQUENCE 597 AA; 65274 MW; 2DAFA8FB7E055851 CRC64;

Query Match 54.7%; Score 354; DB 4; Length 597;
Best Local Similarity 57.0%; Pred. No. 9.1e-31;
Matches 73; Conservative 18; Mismatches 27; Indels 10; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWVQPPGKGLWLGIMWGGNTDYN 60
DB [1]
QY 20 QVQLQWAGLGLKPSQTLISITCTISGFSLSRYSVHWVQPPGKGLWLGIMWGGNTDYN 79
DB [1]
QY 61 SALKSRLSISKDKNQVFLKMNLSITAAADTAVYYCAR-----KGEFYGYDGFVYWGQ 113
DB [1]
QY 80 PSLKSRVTSVDTSKKQLSLKSLSYNAADTAVYYCARVITRASPTDGRYGM-----VWGQ 136
DB [1]
QY 114 GTLVTVSS 121
DB [1]
QY 137 GTTVTVSS 144
DB [1]

RESULT 14
Q96AA6
ID Q96AA6 PRELIMINARY; PRT; 618 AA.
AC Q96AA6;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-WAR-2002 (TRENBLrel. 20, Last annotation update)
Hypothetical 67.8 kDa protein.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017356; AAH17356.1;
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00408; Igc2; 1.
DR SMART; SM00409; Igc1; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
Hypothetical protein; Immunoglobulin domain.
KW PROSITE; PS00290; IG_MHC; UNKNOWN_3.
SEQUENCE 618 AA; 67758 MW; 96BDB4C7C696E0A6 CRC64;

Query Match 54.7%; Score 354; DB 4; Length 618;
Best Local Similarity 57.0%; Pred. No. 9.5e-31;
Matches 73; Conservative 18; Mismatches 27; Indels 10; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWVQPPGKGLWLGIMWGGNTDYN 60
DB [1]
QY 20 QVQLQWAGLGLKPSQTLISITCTISGFSLSRYSVHWVQPPGKGLWLGIMWGGNTDYN 79
DB [1]
QY 61 SALKSRLSISKDKNQVFLKMNLSITAAADTAVYYCAR-----KGEFYGYDGFVYWGQ 113
DB [1]
QY 80 PSLKSRVTSVDTSKKQLSLKSLSYNAADTAVYYCARVITRASPTDGRYGM-----VWGQ 136
DB [1]
QY 114 GTLVTVSS 121
DB [1]
QY 137 GTTVTVSS 144
DB [1]

RESULT 15
Q9UL93
ID Q9UL93 PRELIMINARY; PRT; 116 AA.
AC Q9UL93;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035021; AAD56257.1;
DR HSP; P01772; 2FBA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR NON_TER 1
FT NON_TER 116
FT NON_TER 116
SQ SEQUENCE 116 AA; 12434 MW; 0DA0348154DD6061 CRC64;

Query Match 54.3%; Score 351; DB 4; Length 116;
Best Local Similarity 56.2%; Pred. No. 2.6e-31;
Matches 68; Conservative 20; Mismatches 27; Indels 6; Gaps 2;

QY 2 VOLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWVQPPGKGLWLGIMWGGNTDYN 60
DB [1]
QY 1 VOLVESGGVQVQGRSLRSLSCAASGFTFSSYAMHWVQAPGKGLWVAVISYDGSNKYYA 60
DB [1]
QY 61 SALKSRLSISKDKNQVFLKMNLSITAAADTAVYYCARKEFYGYDGFVYWGQGLTVTVS 120
DB [1]
QY 61 DSVKGRFTISRDNSKNTLYQMNSLRADETAMYYCAGGGGL-----GLGYWGQGLTVTVS 115
DB [1]
QY 121 S 121
DB [1]
QY 116 S 116
DB [1]

RESULT 16
Q9UL75
ID Q9UL75 PRELIMINARY; PRT; 122 AA.
AC Q9UL75;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
```

```
Db 20 QVQLQWAGLGLKPSQTLISITCTISGFSLSRYSVHWVQPPGKGLWLGIMWGGNTDYN 79
QY 61 SALKSRLSISKDKNQVFLKMNLSITAAADTAVYYCAR-----KGEFYGYDGFVYWGQ 113
DB [1]
Db 80 PSLKSRVTSVDTSKKQLSLKSLSYNAADTAVYYCARVITRASPTDGRYGM-----VWGQ 136
QY 114 GTLVTVSS 121
DB [1]
Db 137 GTTVTVSS 144
DB [1]

RESULT 15
Q9UL93
ID Q9UL93 PRELIMINARY; PRT; 116 AA.
AC Q9UL93;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035021; AAD56257.1;
DR HSP; P01772; 2FBA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR NON_TER 1
FT NON_TER 116
FT NON_TER 116
SQ SEQUENCE 116 AA; 12434 MW; 0DA0348154DD6061 CRC64;

Query Match 54.3%; Score 351; DB 4; Length 116;
Best Local Similarity 56.2%; Pred. No. 2.6e-31;
Matches 68; Conservative 20; Mismatches 27; Indels 6; Gaps 2;

QY 2 VOLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWVQPPGKGLWLGIMWGGNTDYN 60
DB [1]
QY 1 VOLVESGGVQVQGRSLRSLSCAASGFTFSSYAMHWVQAPGKGLWVAVISYDGSNKYYA 60
DB [1]
QY 61 SALKSRLSISKDKNQVFLKMNLSITAAADTAVYYCARKEFYGYDGFVYWGQGLTVTVS 120
DB [1]
QY 61 DSVKGRFTISRDNSKNTLYQMNSLRADETAMYYCAGGGGL-----GLGYWGQGLTVTVS 115
DB [1]
QY 121 S 121
DB [1]
QY 116 S 116
DB [1]

RESULT 16
Q9UL75
ID Q9UL75 PRELIMINARY; PRT; 122 AA.
AC Q9UL75;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035039; AAD56275.1; -.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 13719 MW; 56CB0612586A6529 CRC64;

Query Match 54.28; Score 350.5; DB 4; Length 122;
Best Local Similarity 58.48; Pred. No. 3.1e-31;
Matches 73; Conservative 17; Mismatches 28; Indels 7; Gaps 3;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSV--HWVRQPPGKGLWGLMIWGNG-- 56
Db 1 QVQLQSGPGLVKPSQTLSTICTAISGDSVSSNAANWIRQSPRGLWGLGRYYRSKWWY 60
QY 57 TDYNSALKSRLSISKNSKNQVFLKMSLTAAATAVYYCARKGEFYGYGDFYVWGQTL 116
Db 61 NDYRVSVKSRITNPDTSKNQFSLQNSVTPEDTAVYYCARDLELLGQFD--YWGQTL 117
QY 117 VTSS 121
Db 118 VTSS 122

RESULT 17
Q9Y509 PRELIMINARY; PRT; 147 AA.
AC Q9Y509;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Vh3 protein (Fragment).
GN Vh3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=96071149; PubMed=7475288;
RA Cao J., Vescio R.A., Rettig M.B., Hong C.H., Kim A., Lee J.C.,
RA Lichtenstein A.K., Berenson J.R.;
RT "A CD10-positive subset of malignant cells is identified in multiple
RT myeloma using PCR with patient-specific immunoglobulin gene primers.";
RL Leukemia 9:1948-1953(1995).
DR EMBL; S80860; AAD14339.1; -.
DR HSSP; P01772; 2FBA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
SQ SEQUENCE 147 AA; 15768 MW; 8489FCAA7BC925C CRC64;

Query Match 54.28; Score 350.5; DB 4; Length 147;
Best Local Similarity 53.28; Pred. No. 3.9e-31;
Matches 67; Conservative 23; Mismatches 31; Indels 5; Gaps 2;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWGLMIWGNGNDYN 60
Db 1 QVHLVESGGGVQPGKSLRSLSCASGFTTFYGSWVRQAPGKGLDWALISDGSTQYY 60
QY 61 S-ALKSRLSISKNSKNQVFLKMSLTAAATAVYYCARKGEFY----YGDGFYVWGQGT 115

SEQUENCE FROM N.A.
MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035039; AAD56275.1; -.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 13719 MW; 56CB0612586A6529 CRC64;

Query Match 54.28; Score 350.5; DB 4; Length 122;
Best Local Similarity 58.48; Pred. No. 3.1e-31;
Matches 73; Conservative 17; Mismatches 28; Indels 7; Gaps 3;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSV--HWVRQPPGKGLWGLMIWGNG-- 56
Db 1 QVQLQSGPGLVKPSQTLSTICTAISGDSVSSNAANWIRQSPRGLWGLGRYYRSKWWY 60
QY 57 TDYNSALKSRLSISKNSKNQVFLKMSLTAAATAVYYCARKGEFYGYGDFYVWGQTL 116
Db 61 NDYRVSVKSRITNPDTSKNQFSLQNSVTPEDTAVYYCARDLELLGQFD--YWGQTL 117
QY 117 VTSS 121
Db 118 VTSS 122

RESULT 18
Q9BQB8 PRELIMINARY; PRT; 597 AA.
AC Q9BQB8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Unknown (protein for MGC:1905) (protein for MGC:1228).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=MUSCLE;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006180; AAH06180.1; -.
DR EMBL; BC001872; AAH01872.1; -.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 4.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
SQ SEQUENCE 597 AA; 65300 MW; 2DAFAD50A6375851 CRC64;

Query Match 54.18; Score 350; DB 4; Length 597;
Best Local Similarity 57.08; Pred. No. 2.9e-30;
Matches 73; Conservative 17; Mismatches 28; Indels 10; Gaps 2;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWGLMIWGNGNDYN 60
Db 20 QVQLQSGAGLLKPSQTLSTCTGCGYGSFSGYGSWVRQPPGKGLWGLMIWGNGNDYN 79
QY 61 SALKSRLSISKNSKNQVFLKMSLTAAATAVYYCAR-----KGEFYGYGDFYVWGQ 113
Db 80 PSLKSRVTISVDTSKKQLSLKSSVNAADTAVYYCARVITRASPGTDGRYGM---VWGQ 136
QY 114 GTLVTVSS 121
Db 137 GTTVTVSS 144

RESULT 19
Q96BB9 PRELIMINARY; PRT; 597 AA.
AC Q96BB9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 65.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=B-CCELL;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015760; AAH15760.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_5.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 597 AA; 65039 MW; 4FCA3AD8E263D9 CRC64;

Query Match 54.1%; Score 350; DB 4; Length 597;
Best Local Similarity 55.2%; Pred. No. 2.5e-30;
Matches 69; Conservative 21; Mismatches 31; Indels 4; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLITCTISGFSLSRYSVHVRQPPGKGLWGLMIGWGNTDY 59
20 EVQLLESGGGLVQPGSLRLSCAASGFSFSSYAMNVRQAPGKGLWWSAISGSGSTYY 79
60 NSALKSRSLISKDNKNOVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVWGQGLT 116
80 ADSVKGRTISRDNSTLTLYLQMNLSRAEDTAVYCAKDPGRGYSASGNVTREDYWGQGLT 139
117 VTVSS 121
140 VTVSS 144
DB [1]

RESULT 20
Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024405; AAH24405.1; -.
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match 53.6%; Score 346.5; DB 11; Length 469;
Best Local Similarity 51.6%; Pred. No. 4.6e-30;
Matches 65; Conservative 25; Mismatches 31; Indels 5; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLITCTISGFSLSRYSVHVRQPPGKGLWGLMIGWGNTDY 57
20 EVNLVESGGGLVQPGSLRLSCAASGFTTIDYMSVRQPPGKALEWGLFIRNKANGYTT 79
58 DYNALKSRSLISKDNKNOVFLKMNLSLTAAADTAVYCA--RKGEFYGYDGFVWGQGT 115
80 EYASVKGRTISRDNSTLYLQMNLSRAEDTAVYCAKDPGRGYSASGNVTREDYWGQGLT 139
116 LTVSS 121
140 LTVSS 145
DB [1]

RESULT 21
Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical 52.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025985; AAH25985.1; -.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 53.4%; Score 345.5; DB 4; Length 473;
Best Local Similarity 57.4%; Pred. No. 6e-30;
Matches 70; Conservative 22; Mismatches 25; Indels 5; Gaps 3;

QY 1 QVQLKESGPGLVKPSQTLITCTISGFSLS--SRYSVHVRQPPGKGLWGLMIGWGNTD 58
27 RULQESGPGLLKPSYTLSTCTVSGDSVASSYVWGVQPPGKGLWIGTINFSGNMY 86
59 YNSALKSRSLISKDNKNOVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVWGQGLT 118
87 YPSLSRSRVMTSADMSSENFYKLDSTVTAADTAVYCA--AGHLVWVGFG--AHWGQGLVS 143
119 VS 120
144 VS 145
DB [1]

RESULT 22
Q9UL90 PRELIMINARY; PRT; 113 AA.
AC Q9UL90;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region (Fragment).
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035024; AAD56260.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; Igv; 1.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;

Query Match 53.2%; Score 344; DB 4; Length 113;
Best Local Similarity 54.9%; Pred. No. 1.5e-30;
Matches 67; Conservative 18; Mismatches 27; Indels 10; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLITCTISGFSLSRYSVHVRQPPGKGLWGLMIGWGNTDY 59
20 EVQLVESGGGVQPGSLRLSCAASGFTTSSYGMHVRQAPGKGLWVAFIRYDGSNKYY 60
60 NSALKSRSLISKDNKNOVFLKMNLSLTAAADTAVYVCARKGEFYGYDGFVWGQGLT 119
```

```
DB 61 ADSVKGRTISRDNKNTLYLQMNLSRAEDTAVYYCAR-----DLNWGQGLTV 111
QY 120 SS 121
DB 112 SS 113

RESULT 23
Q9HCC1 PRELIMINARY; PRT; 112 AA.
AC Q9HCC1
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Single chain Fv (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RA Kikuchi M., Takeda C., Tsujimoto Y., Asada S., Nagata K.;
RT "An antibody fragment2A3 specific for native lysozyme: Isolation from a
human synthetic phage display library and characterization."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049915; BAB16829.1; -.
DR HSP; P01772; 2FB4.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
FT NON_TER 1
FT NON_TER 112
FT NON_TER 112
SQ SEQUENCE 112 AA; 12243 MW; 24F1A45EC3B84788 CRC64;

Query Match 53.1%; Score 343.5; DB 4; Length 112;
Best Local Similarity 55.9%; Pred. No. 1.7e-30;
Matches 66; Conservative 19; Mismatches 26; Indels 7; Gaps 3;

QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDY 59
1 EVQLVESGGGVVPRPGSLRISCAASGFTDDYGMISWVRQAPGKGLWVSGINWNGSGTGY 60
60 NSALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGLTV 117
61 ADSVKGRTISRDNKNTLYLQMNLSRAEDTAVYYCARR---RYALD---YWGQGLTV 112

DB 61 ADSVKGRTISRDNKNTLYLQMNLSRAEDTAVYYCARR---RYALD---YWGQGLTV 112

RESULT 24
Q9UL72 PRELIMINARY; PRT; 118 AA.
AC Q9UL72
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
(Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=98271139; PubMed=9614934;
RA Wu X., Liu B., Van der Merve P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus.";
```

```
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035042; AAD56278.1; -.
DR HSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944B2D5CCA CRC64;

Query Match 52.2%; Score 337.5; DB 4; Length 118;
Best Local Similarity 55.3%; Pred. No. 8.3e-30;
Matches 68; Conservative 21; Mismatches 27; Indels 7; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDY 60
1 EVQLVESGGGVVPRPGSLRISCAASGFTVSSNMWVRQAPGKGLSVSVYSGGSYYA 60
61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARK--GEFYGYDGFVYWGQGLTV 118
61 DSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCARDPFGEL-----FDVWGQGLTV 115

QY 119 VSS 121
DB 116 VSS 118

RESULT 25
Q920E7 PRELIMINARY; PRT; 119 AA.
AC Q920E7
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Pterin-mimicking anti-idiotope heavy chain variable region
(Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RA Atkin J.D., Iape A., Jennings I.G., Horaitis O., Cotton R.G.H.;
RT "Definition of the idiotope of Pterin-Mimicking Antibodies Expressed
in Mammalian Cells."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF307937; AAL09421.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
FT NON_TER 1
FT NON_TER 119
SQ SEQUENCE 119 AA; 13025 MW; F6E904044381CA7C CRC64;

Query Match 52.1%; Score 337; DB 11; Length 119;
Best Local Similarity 54.5%; Pred. No. 9.6e-30;
Matches 67; Conservative 21; Mismatches 29; Indels 6; Gaps 3;

QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGN-TDY 59
1 EVQLVESGGDLVKPGGSLKLSCAASGFTFSSYGMWVRQTPDKRLEWVATISGGSYYV 60
60 NSALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKGEFYGYD-GFVYWGQGLTV 118
61 PDSVKGRTISRDNKNTLYLQMNLSKSDTAVYYCARGD----YDVGFAIWGQGLTV 116

QY 119 VSS 121
DB 117 VSA 119

Search completed: June 23, 2003, 14:03:31
Job time : 16.6613 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:00:06 : Search time 7.07979 Seconds
(without alignments)
502.864 Million cell updates/sec

Title: US-10-056-052A-20

Perfect score: 647

Sequence: 1 QVQLKESGGLVKPSQTLISI.....YYGDGFVYWGQGLTVTVSS 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database :

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	506	78.2	119	1	US-07-634-278-60
2	506	78.2	119	1	US-08-477-728-60
3	506	78.2	119	1	US-08-474-040-60
4	506	78.2	119	1	US-08-487-200-60
5	506	78.2	119	4	US-08-484-537-60
6	506	78.2	138	1	US-07-634-278-33
7	506	78.2	138	1	US-08-477-728-33
8	506	78.2	138	1	US-08-474-040-33
9	506	78.2	138	1	US-08-487-200-33
10	506	78.2	138	4	US-08-484-537-33
11	502	77.6	113	2	US-08-606-293-6
12	494.5	76.4	137	2	US-08-621-751A-8
13	492	76.0	119	2	US-08-752-844-16
14	492	76.0	119	2	US-08-591-196-16
15	491.5	76.0	242	6	5455030-15
16	487.5	75.3	120	4	US-08-652-558-8
17	487.5	75.3	120	4	US-09-254-189-5
18	487.5	75.3	239	2	US-08-860-174A-2
19	487	75.3	113	2	US-08-606-293-2
20	486.5	75.2	120	2	US-08-652-558-38
21	486	75.1	222	2	US-08-190-199A-67
22	486	75.1	235	2	US-08-190-199A-61
23	481.5	74.4	137	2	US-08-621-751A-4
24	479.5	74.1	120	2	US-08-652-558-7
25	479.5	74.1	120	4	US-09-254-189-4
26	471.5	72.9	116	3	US-08-397-411-3
27	471.5	72.9	121	3	US-08-881-037-67

28	471.5	72.9	273	3	US-08-397-411-6	Sequence 6, Appl
29	471.5	72.9	446	3	US-08-397-411-7	Sequence 7, Appl
30	468	72.3	116	2	US-08-308-494A-21	Sequence 21, Appl
31	467	72.2	119	1	US-08-667-769A-15	Sequence 15, Appl
32	467	72.2	119	1	US-08-667-769A-62	Sequence 62, Appl
33	467	72.2	119	5	PCT-US95-17082A-15	Sequence 15, Appl
34	467	72.2	119	5	PCT-US95-17082A-62	Sequence 62, Appl
35	466.5	72.1	120	2	US-08-652-558-5	Sequence 5, Appl
36	466.5	72.1	120	2	US-08-652-558-2	Sequence 2, Appl
37	466.5	72.1	120	4	US-09-254-189-2	Sequence 4, Appl
38	466.5	72.1	140	4	US-08-943-136-4	Sequence 4, Appl
39	466.5	72.1	140	4	US-08-973-518-4	Sequence 4, Appl
40	465.5	71.9	239	6	5455030-13	Patent No. 5455030
41	464.5	71.8	120	4	US-08-652-558-6	Sequence 6, Appl
42	464.5	71.8	120	4	US-09-254-189-3	Sequence 3, Appl
43	463	71.6	119	1	US-08-467-420A-62	Sequence 62, Appl
44	463	71.6	119	1	US-08-470-110A-62	Sequence 62, Appl
45	463	71.6	119	2	US-08-940-371-62	Sequence 62, Appl
46	459.5	71.0	120	2	US-08-652-558-9	Sequence 9, Appl
47	459.5	71.0	120	4	US-09-254-189-6	Sequence 6, Appl
48	458.5	70.9	120	4	US-08-057-430A-26	Sequence 26, Appl
49	458	70.8	152	2	US-08-752-844-4	Sequence 4, Appl
50	458	70.8	152	2	US-08-591-196-4	Sequence 4, Appl
51	458	70.8	152	4	US-09-192-838B-4	Sequence 4, Appl
52	458	70.8	263	2	US-08-752-844-66	Sequence 66, Appl
53	455.5	70.4	107	1	US-07-942-245-14	Sequence 14, Appl
54	455.5	70.4	120	2	US-08-353-372A-36	Sequence 36, Appl
55	455	70.3	119	3	US-08-483-749A-14	Sequence 14, Appl
56	454.5	70.2	112	4	US-09-189-129-3	Sequence 3, Appl
57	452.5	69.9	116	3	US-08-397-411-4	Sequence 4, Appl
58	449	69.4	264	4	US-08-564-164A-4	Sequence 4, Appl
59	447.5	69.2	111	2	US-08-470-139-6	Sequence 6, Appl
60	447.5	69.2	111	4	US-09-347-061-6	Sequence 6, Appl
61	447.5	69.2	111	4	US-09-537-911A-67	Sequence 67, Appl
62	447	69.1	119	1	US-07-634-278-37	Sequence 37, Appl
63	447	69.1	119	1	US-07-634-278-61	Sequence 61, Appl
64	447	69.1	119	1	US-08-477-728-37	Sequence 37, Appl
65	447	69.1	119	1	US-08-477-728-61	Sequence 61, Appl
66	447	69.1	119	1	US-08-474-040-37	Sequence 37, Appl
67	447	69.1	119	1	US-08-474-040-61	Sequence 61, Appl
68	447	69.1	119	1	US-08-487-200-37	Sequence 37, Appl
69	447	69.1	119	1	US-08-487-200-61	Sequence 61, Appl
70	447	69.1	119	4	US-08-484-537-37	Sequence 37, Appl
71	447	69.1	119	4	US-08-484-537-61	Sequence 61, Appl
72	437	67.5	97	3	US-08-881-037-66	Sequence 66, Appl
73	434	67.1	115	3	US-08-881-037-22	Sequence 22, Appl
74	431	66.6	119	4	US-09-025-769B-39	Sequence 39, Appl
75	431	66.6	119	4	US-09-025-769B-65	Sequence 65, Appl
76	429	66.3	111	1	US-08-467-420A-15	Sequence 15, Appl
77	429	66.3	111	1	US-08-470-110A-15	Sequence 15, Appl
78	429	66.3	111	2	US-08-940-371-15	Sequence 15, Appl
79	429	66.3	111	3	US-08-637-647-15	Sequence 15, Appl
80	427.5	66.1	215	4	US-09-170-769A-6	Sequence 6, Appl
81	426	65.8	119	1	US-08-467-420A-19	Sequence 19, Appl
82	426	65.8	119	1	US-08-470-110A-19	Sequence 19, Appl
83	426	65.8	119	1	US-08-667-769A-19	Sequence 19, Appl
84	426	65.8	119	2	US-08-940-371-19	Sequence 19, Appl
85	426	65.8	119	3	US-08-637-647-19	Sequence 19, Appl
86	426	65.8	119	5	PCT-US95-17082A-19	Sequence 19, Appl
87	423	65.4	107	1	US-08-122-546-12	Sequence 12, Appl
88	423	65.4	107	2	US-08-764-938-12	Sequence 12, Appl
89	423	65.4	107	3	US-09-131-052-12	Sequence 12, Appl
90	423	65.4	107	4	US-09-131-053A-12	Sequence 12, Appl
91	421.5	65.1	144	2	US-08-116-778B-36	Sequence 36, Appl
92	421.5	65.1	144	2	US-08-438-562-36	Sequence 36, Appl
93	421.5	65.1	144	2	US-08-483-528B-100	Sequence 100, Appl
94	420.5	65.0	118	4	US-09-025-769B-25	Sequence 25, Appl
95	419.5	64.8	135	2	US-08-470-139-28	Sequence 28, Appl
96	419.5	64.8	135	4	US-09-347-061-28	Sequence 28, Appl
97	411.5	63.6	122	1	US-08-360-125-11	Sequence 11, Appl
98	411.5	63.6	122	2	US-08-450-578-11	Sequence 11, Appl
99	411.5	63.6	122	2	US-09-017-628-11	Sequence 11, Appl
100	411.5	63.6	122	2	US-09-014-880-11	Sequence 11, Appl

Db 119 A 119

RESULT 2
US-08-477-728-60
: Sequence 60, Application US/0847728
: Patent No. 5585089
: GENERAL INFORMATION:
: APPLICANT: QUEEN, Cary L.
: APPLICANT: SCHNEIDER, William P.
: APPLICANT: SELICK, Harold E.
: TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
: NUMBER OF SEQUENCES: 113
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Crew LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: Palo Alto
: STATE: California
: COUNTRY: US
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/477,728
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,728
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/634,278
 FILING DATE: 10 DEC 1990

FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310 353

AFFILIATION NUMBER: US 01/310,232
 FILING DATE: 13-FEB-1989
 PRIOR AFFILIATION DATE:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975

FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:

NAME: Smith, William M
REGISTRATION NUMBER: 30.223

REGISTRATION NUMBER: 50,225
REFERENCE/DOCKET NUMBER: 11823-001
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 326-2400

TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids

SERGIN. 119 amino acids
 TYPE: amino acid.
 STRANDEFENSE. single

STRADEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
-477-728-60

Match	78 %	Score
50		

Method	Local Similarity	Pred. No.	Score
Local Similarity	75.2%	75.2%	78.2%
Pred. No.	75.2%	75.2%	78.2%
Score	78.2%	78.2%	78.2%

ches 91; Conservative 16; Mism

1 QVQLKESGPGLVKPSQTLSTITCTISGF9

1 QVQLKQSGPGLVQPSSQLSITCTVSGFS

61 SALKSRLISKDNSKNQVFLKMNSLTAF

61 AAFISRLTISKDNSKSQVFFKVNSLQPA

121 S 121

119 Å 119

0
1
1
2
2
3

RESULT 3

US-08-474-040-60
: Sequence 60, Application US/08474040
: Patent No. 5693761
: GENERAL INFORMATION:
: APPLICANT: QUEEN, Cary L.
: APPLICANT: CO, Man Sung
: APPLICANT: SCHNEIDER, William P.
: APPLICANT: LANDOLET, Nicholas F.
: APPLICANT: COELINGH, Kathleen L.
: APPLICANT: SELICK, Harold E.
: TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
: NUMBER OF SEQUENCES: 113
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend Kourie and Crew
: STREET: 379 Lytton Avenue
: CITY: Palo Alto
: STATE: California
: COUNTRY: US
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474,040
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/634,278
: FILING DATE: 19-DEC-1990
: APPLICATION NUMBER: US 07/590,274
: FILING DATE: 28-SEP-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/310,252
: FILING DATE: 13-FEB-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/290,975
: FILING DATE: 28-DEC-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Smith, William M
: REGISTRATION NUMBER: 30,223
: REFERENCE/DOCKET NUMBER: 11823-002600
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 60:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 119 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-474-040-60

Query Match 78.2%; Score 506; DB 1; Length 119;
Best Local Similarity 75.2%; Pred. No. 4.5e-44;
Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;
Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWLGMIWGSGNTDYN 60
Db 1 QVQLKQSGPGLVQPSQSLISITCTVSGFSVTSYGVHWIRQSPGKGLWLGVTWWSGSDYN 60
Qy 61 SALKSRISISKDKNQVFLKMSLTADTAATVYYCARKGEFYGYDGFVYWGQGLTVTS 120
Db 61 AAFISRLTISKDNSKQVFFKVNLSQPADTAIYYCARAGD--YNYDGFAYWGQGLTVTS 118
Qy 121 S 121
Db 119 A 119

RESULT 4

US-08-487-200-60
: Sequence 60, Application US/08487200
: Patent No. 5693762
: GENERAL INFORMATION:
: APPLICANT: QUEEN, Cary L.
: APPLICANT: CO, Man Sung
: APPLICANT: SCHNEIDER, William P.
: APPLICANT: LANDOLET, Nicholas F.
: APPLICANT: COELINGH, Kathleen L.
: APPLICANT: SELICK, Harold E.
: TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
: NUMBER OF SEQUENCES: 113
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew
: STREET: 379 Lytton Avenue
: CITY: Palo Alto
: STATE: California
: COUNTRY: US
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487,200
: FILING DATE: 7-JUN-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/634,278
: FILING DATE: 19-DEC-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/590,274
: FILING DATE: 28-SEP-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/310,252
: FILING DATE: 13-FEB-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/290,975
: FILING DATE: 28-DEC-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Smith, William M
: REGISTRATION NUMBER: 30,223
: REFERENCE/DOCKET NUMBER: 11823-002610
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 60:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 119 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-487-200-60

Query Match 78.2%; Score 506; DB 1; Length 119;
Best Local Similarity 75.2%; Pred. No. 4.5e-44;
Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;
Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWLGMIWGSGNTDYN 60
Db 1 QVQLKQSGPGLVQPSQSLISITCTVSGFSVTSYGVHWIRQSPGKGLWLGVTWWSGSDYN 60
Qy 61 SALKSRISISKDKNQVFLKMSLTADTAATVYYCARKGEFYGYDGFVYWGQGLTVTS 120
Db 61 AAFISRLTISKDNSKQVFFKVNLSQPADTAIYYCARAGD--YNYDGFAYWGQGLTVTS 118
Qy 121 S 121
Db 119 A 119

RESULT 5
 US-08-484-537-60
 ; Sequence 60, Application US/08484537
 ; Patent No. 6180370
 ; GENERAL INFORMATION:
 ; APPLICANT: QUEEN, Cary L.
 ; APPLICANT: CO, Man Sung
 ; APPLICANT: SCHNEIDER, William P.
 ; APPLICANT: LANDOLFI, Nicholas F.
 ; APPLICANT: COELINGH, Kathleen L.
 ; APPLICANT: SELICK, Harold E.
 ; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
 ; NUMBER OF SEQUENCES: 113
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Kourie and Crew
 ; STREET: 379 Lytton Avenue
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 94301
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,537
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/634,278
 ; FILING DATE: 19-DEC-1990
 ; APPLICATION NUMBER: US/07/590,274
 ; FILING DATE: 28-SEP-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/310,252
 ; FILING DATE: 13-FEB-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/290,975
 ; FILING DATE: 28-DEC-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith, William M
 ; REGISTRATION NUMBER: 30,223
 ; REFERENCE/DOCKET NUMBER: 11823-002600
 ; TELEPHONE: (415) 326-2400
 ; TELEFAX: (415) 326-2422
 ; INFORMATION FOR SEQ ID NO: 60:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 119 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-484-537-60

[illegible]

```

RESULT 6
US-07-634-278-33
; Sequence 33, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,278
; FILING DATE: 19-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-634-278-33

```

Query Match	78.2%;	Score 506;	DB 1;	Length 138;
Best Local Similarity	75.2%;	Pred. No. 5.4e-44;		
Matches	91; Conservative 16;	Mismatches 12;	Indels 2;	Gaps 1;
QY	1	OVLKESGCLVKPQSOTLSITCTISGFSLRSRYSVHWVROP	PGKLEWLGNICGGNTDYN	60
Dd	20	OVLKOSGGLVPQSLSITCVSGFSVTSGVHWIROS	PGKLEWLGVINSGGSTDYN	79
QY	61	SALKSRSLTISKNSKNQVFLKMNLSLTAADTAVTYVCARKGEFY	YGDFVFWGGTLTVTS	120
Dd	80	AAFISRLTISKNSKSQVFVKVNSLPADTAIIYCARGD--	INYDCFAWVGOGTLTVTS	137
QY	121	S	121	
Dd	138	A	138	

RESULT 7

APPLICANT: QUEEN, Cary L.
 APPLICANT: CO, Man Sung
 APPLICANT: SCHNEIDER, William P.
 APPLICANT: LANDOLFI, Nicholas F.
 APPLICANT: COELINGH, Kathleen L.
 APPLICANT: SELICK, Harold E.
 TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBLINS
 NUMBER OF SEQUENCES: 113
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew
 STREET: 379 Lytton Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: US
 ZIP: 94301
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,200
 FILING DATE: 7-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/634,278
 FILING DATE: 19-DEC-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/590,274
 FILING DATE: 28-SEP-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/310,252
 FILING DATE: 13-FEB-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/290,975
 FILING DATE: 28-DEC-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, William M
 REGISTRATION NUMBER: 30,223
 REFERENCE/DOCKET NUMBER: 11823-002610
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 326-2400
 TELEFAX: (415) 326-2422
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 138 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-487-200-33

Query Match 78.2%; Score 506; DB 1; Length 138;
 Best Local Similarity 75.2%; Pred. No. 5.4e-44;
 Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;
 Qy 1 QVQLKESGPGLVKPSQTLSTICTSGFSLRSYVHWVRQPPGKGLWGLMIGGGNTDYN 60
 Db 20 QVQLKQSGPGLVQPSQSLSTICTSGFSLRSYVHWVRQPPGKGLWGLVWSGGSTDYN 79
 Qy 61 SALKSRISISKDKNQVFLKMNLSLTAAATVAVYCARKEFYGYDGFVYWGQGLTVTS 120
 Db 80 AAFISRLTISKDNKSQVEFKVNSLQPADTAIYYCARAGD--YNYDGFAYWGQGLTVTS 137
 Qy 121 S 121
 Db 138 A 138

RESULT 10
 US-08-484-537-33
 ; Sequence 33, Application US/08484537
 ; Patent No. 6180370
 ; GENERAL INFORMATION:

APPLICANT: QUEEN, Cary L.
 APPLICANT: CO, Man Sung
 APPLICANT: SCHNEIDER, William P.
 APPLICANT: LANDOLFI, Nicholas F.
 APPLICANT: COELINGH, Kathleen L.
 APPLICANT: SELICK, Harold E.
 TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBLINS
 NUMBER OF SEQUENCES: 113
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew
 STREET: 379 Lytton Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: US
 ZIP: 94301
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,537
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/634,278
 FILING DATE: 19-DEC-1990
 APPLICATION NUMBER: US 07/590,274
 FILING DATE: 28-SEP-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/310,252
 FILING DATE: 13-FEB-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/290,975
 FILING DATE: 28-DEC-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, William M
 REGISTRATION NUMBER: 30,223
 REFERENCE/DOCKET NUMBER: 11823-002600
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 326-2400
 TELEFAX: (415) 326-2422
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 138 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-484-537-33

Query Match 78.2%; Score 506; DB 4; Length 138;
 Best Local Similarity 75.2%; Pred. No. 5.4e-44;
 Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;
 Qy 1 QVQLKESGPGLVKPSQTLSTICTSGFSLRSYVHWVRQPPGKGLWGLMIGGGNTDYN 60
 Db 20 QVQLKQSGPGLVQPSQSLSTICTSGFSLRSYVHWVRQPPGKGLWGLVWSGGSTDYN 79
 Qy 61 SALKSRISISKDKNQVFLKMNLSLTAAATVAVYCARKEFYGYDGFVYWGQGLTVTS 120
 Db 80 AAFISRLTISKDNKSQVEFKVNSLQPADTAIYYCARAGD--YNYDGFAYWGQGLTVTS 137
 Qy 121 S 121
 Db 138 A 138

RESULT 11
 US-08-606-293-6
 ; Sequence 6, Application US/08606293
 ; Patent No. 5874082
 ; GENERAL INFORMATION:
 ; APPLICANT: de Boer, Mark

TITLE OF INVENTION: Humanized Anti-CD40 Monoclonal Antibodies and
TITLE OF INVENTION: Fragments Capable of Blocking B Cell Activation
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION Intellectual Property - R440
STREET: 4560 Horton Street, P.O. Box 8097
CITY: Emeryville
STATE: California
COUNTRY: United States of America
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606.293
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/070.158
FILING DATE: 28-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Savereide, Paul B.
REGISTRATION NUMBER: 36,914
REFERENCE/DOCKET NUMBER: 27527/33157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2718
TELEFAX: (510) 655-3542
TELEX: n/a
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-606-293-6

Query Match 77.6%; Score 502; DB 2; Length 113;
Best Local Similarity 81.0%; Pred. No. 1.le-43;
Matches 98; Conservative 7; Mismatches 8; Indels 8; Gaps 2;
Qy 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRQPPGKGLWLGIMWGNTDYN 60
Db 1 QVQLVESGPGLVKPSQSLSTCTISGFSLSRYSVHVVRRQPPGKGLWLGIMWGNTDYN 60
Qy 61 SALKSRLSISKDNKNOVFLKMNLSLTAAADTAVVYCARKEFYCYDGFVYWGQTLTVTS 120
Db 61 SALKSRLTISKDTSKNOVFLKMNLSRAEDTAMVYCVRT-----DG-DYWGQGTTLTVTS 112
Qy 121 S 121
Db 113 S 113

RESULT 12
US-08-621-751A-8
Sequence 8, Application US/08621751A
Patent No. 5882644
GENERAL INFORMATION:
APPLICANT: Chang, Chung N.
APPLICANT: Landolfi, Nicholas F.
APPLICANT: Martin, Ulrich
TITLE OF INVENTION: MONOCLONAL ANTIBODIES SPECIFIC FOR THE
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR BETA RECEPTOR AND
TITLE OF INVENTION: METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER LLP
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CA
COUNTRY: USA

ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/621.751A
FILING DATE: 22-MAR-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lehnhardt, Susan K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 321152000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
TELEX: 706141 MRSN FOERS SFO
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-621-751A-8

Query Match 76.4%; Score 494.5; DB 2; Length 137;
Best Local Similarity 76.9%; Pred. No. 7.7e-43;
Matches 93; Conservative 14; Mismatches 11; Indels 3; Gaps 1;
Qy 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRQPPGKGLWLGIMWGNTDYN 60
Db 20 QVQLQESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRQPPGKGLWLGIMWGNTDYN 79
Qy 61 SALKSRLSISKDNKNOVFLKMNLSLTAAADTAVVYCARKEFYCYDGFVYWGQTLTVTS 120
Db 80 SALKSRLTISKDTSKNOVFLKMNLSLTAAADTAVVYCARKEFYCYDGFVYWGQTLTVTS 136
Qy 121 S 121
Db 137 S 137

RESULT 13
US-08-752-844-16
Sequence 16, Application US/08752844
Patent No. 5935821
GENERAL INFORMATION:
APPLICANT: Chatterjee, Malaya
APPLICANT: Foon, Kenneth A.
APPLICANT: Chatterjee, Sunil K.
TITLE OF INVENTION: MONOCLONAL ANTIBODY IA7 AND USE FOR THE
TITLE OF INVENTION: TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752.844
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schliff, J. Michael
REGISTRATION NUMBER: 40,253

US-08-591-196-16						
		Query Match	76.0%;	Score 492;	DB 2:	Length 119;
		Best Local Similarity	78.7%;	Pred. No. 1.2e-42;		
		Matches 96;	Conservative 8;	Mismatches 14;	Indels 4;	Gaps 2
Qy	1	QVQLKESGPGLVKPSQTLSITCTISGTSGSLRSYSHVHWVRQPPGKGLEWLMINGGGNTDYN	60			
Dd	1	QVQLKESGPGLVAPSQSLSTICTVTSGFSLTSYGVHWRQPPGKGLEWLVGINIGSDTNYN	60			
Qy	61	SALKSRLSISKDNSKNQVFLLKMNSLTAADTAIVYTCAEK-GEFYTYGDGFVWGQGLTVTV	119			
Dd	61	SALKSPSTISKNSKSQFVKMNSLQMDTPAYVCABEYVVVVAVMD---VSCQCTSVTV	117			

Qy 120 SS 121
11
Db 118 SS 119

RESULT 15
5455030-15
; Patent No. 5455030
; APPLICANT : LADNER, ROBERT C.; BIRD, ROBERT E.; HARDMAN, KARL
; TITLE OF INVENTION : IMMUNOTHERAPY USING SINGLE CHAIN
; POLYPEPTIDE BINDING MOLECULES
; NUMBER OF SEQUENCES : 24
; CURRENT APPLICATION DATA :
; APPLICATION NUMBER : US/08/40,440
; FILING DATE : 1-APP-1993

```

; APPLICATION NUMBER: 512,910
; FILING DATE: 25-APR-1990
; APPLICATION NUMBER: 299,617
; FILING DATE: 19-JAN-1989
; APPLICATION NUMBER: 92,110
; FILING DATE: 02-SEP-1987
; APPLICATION NUMBER: 902,971
; FILING DATE: 01-SEP-1986
; SEQ ID NO:15:
; LENGTH: 242
5455030-15

Query Match 76.0%; Score 491.5; DB 6; Length 242;
Best Local Similarity 77.9%; Pred. No. 3e-42;
Matches 95; Conservative 9; Mismatches 13; Indels 5; Gaps 2

Qy 2 VOLKESGPGLVKPQSOTLSITCTISGFSLSRYSVHVRQPPGKGLEWLGMIWGGNTDYS 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 VOLKESGPGLVAPQSOSTSCTCTVSCTSLTAWCVHVRQPPGKCFEIGCVLHCCNTDYS 182
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

124 VOLKESGPVLVAPSQSLISITCTVSGFSLTNGVHVWRQPPGKGLEWLGVIWAGGNNTYNS 180
62 ALKSRLSISKDMSKNVOFLKMNSLTAADTAIVYYCARKGE--FYGYDCGFVYGOGTLVTV 119
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||
184 ALMSRLSISKDMSKSQVFLKMNSLOIDDTAIIYCARKLERIFYYAMD---YWGGTSVTV 240
   |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||
QY      120 SS 121
       ||
Db      241 SS 242

```

RESULT 16
US-08-652-558-8
; Sequence 8, Application US/08652558
; Patent No. 5861155
; GENERAL INFORMATION:
; APPLICANT: LIN, AUGUSTINE YEE-THARN
; TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & WITCOFF
; STREET: 75 STATE STREET, 23RD FLOOR
; CITY: BOSTON
; STATE: MASSACHUSETTS

COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
FILING DATE: JUNE 6, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB94/00387
FILING DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKWICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-8

Query Match 75.3%; Score 487.5; DB 2; Length 120;
Best Local Similarity 76.4%; Pred. No. 3.4e-42;
Matches 94; Conservative 12; Mismatches 12; Indels 5; Gaps 2;
Qy 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGSGNTDYN 60
Db 1 QVQLQESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGSGNTDYN 60
Qy 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARK--GEFYGYDGFVYWGQGLT 118
Db 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARK--GEFYGYDGFVYWGQGLT 118
Qy 119 VSS 121
Db 118 VSS 120

RESULT 17
US-09-254-189-5
Sequence 5, Application US/09254189
Patent No. 6150792
GENERAL INFORMATION:
APPLICANT: Lundquist, Tomas
TITLE OF INVENTION: Sequence Listing
Patent No. 6150792
FILE REFERENCE: 3526/00000
CURRENT APPLICATION NUMBER: US/09/254,189
CURRENT FILING DATE: 1999-03-02
EARLIER APPLICATION NUMBER: 00/000,000
EARLIER FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 5
LENGTH: 120
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:humanized
US-09-254-189-5

Query Match 75.3%; Score 487.5; DB 4; Length 120;
Best Local Similarity 76.4%; Pred. No. 3.4e-42;
Matches 94; Conservative 12; Mismatches 12; Indels 5; Gaps 2;

Qy 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGSGNTDYN 60
Db 1 QVQLQESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGSGNTDYN 60
Qy 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARK--GEFYGYDGFVYWGQGLT 118
Db 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARK--GEFYGYDGFVYWGQGLT 118
Qy 119 VSS 121
Db 118 VSS 120

RESULT 18
US-08-860-174A-2
Sequence 2, Application US/08860174A
Patent No. 5989830
GENERAL INFORMATION:
APPLICANT: DAVIS, Paul James
APPLICANT: VAN DER LOGT, Cornelis Paul Erik
APPLICANT: VERHOEIJEN, Martine Elisa
APPLICANT: WILSON, Steve
TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP
STREET: 1100 New York Avenue, N.W.
CITY: WASHINGTON, D.C.
STATE:
COUNTRY: UNITED STATES
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS .DOS Text
SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/860,174A
FILING DATE: June 16, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95307332.7
FILING DATE: October 16, 1995
APPLICATION NUMBER: PCT/EP/96/03605
FILING DATE: August 14, 1996
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-860-174A-2

Query Match 75.3%; Score 487.5; DB 2; Length 239;
Best Local Similarity 79.3%; Pred. No. 7.5e-42;
Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;
Qy 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGSGNTDYN 60
Db 124 QVQLQESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGSGNTDYN 183
Qy 61 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKGEFYCYDGFVYWGQGLT 120
Db 184 SALKSRLSISKDNKSNQVFLKMNLSLTAAADTAVYYCARKGEFYCYDGFVYWGQGLT 238
Qy 121 S 121
Db 239 S 239
RESULT 19

RESULT 20
US-08-652-558-38
: Sequence 38, Application US/08632558
: Patent No. 5861155
: GENERAL INFORMATION:
: APPLICANT: LIN, AUGUSTINE YEE-THARN
: TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
: TITLE OF INVENTION: THEREOF
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BANNER & WITCOFF
: STREET: 75 STATE STREET, 23RD FLOOR
: CITY: BOSTON

```

Query Match      75.3%; Score 487; DB 2; Length 113;
Best Local Similarity 78.5%; Pred. No. 3.5e-42;
Matches 95; Conservative 9; Mismatches 9; Indels 8; Gaps 2;

1 QVQLKESGPGVLVPSQTLSTCTISGFSLSRYSVHHVRPPGKGLWLGIMGGNTDYN 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 QVQLKEGPGVLVAPSQSLSTCTVSGFSLSRYSVYHWRPFGKGLWLGIMMGGSTDYN 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

61 SALKSRLSISKDNSKNQVFLKMNLSAAATPAVYICARKGEFYGYDGFVYWGQGLTVTS 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

61 SALKSRLSISKDTSKQVFLKMNLSLQTDTPAMYCYVRT-----DG-DYWGQGLTSVTS 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

121 S 121
|
db 113 S 113
db

```

RESULT 22
US-08-190-199A-61
; Sequence 61, Application US/08190199A
; Patent No. 5830663
; GENERAL INFORMATION:
; APPLICANT: EMBLETON, Michael J.
; APPLICANT: GOROCHOV, Guy
; APPLICANT: JONES, Peter T.
; APPLICANT: WINTER, Gregory P.
; TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190.199A
; FILING DATE: 13-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/01483
; FILING DATE: 10-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9212419.7
; FILING DATE: 11-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9117352.6
; FILING DATE: 10-AUG-1991
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:

```

Query Match      74.4%; Score 481.5; DB 2; Length 137;
Best Local Similarity 76.0%; Pred. No. 1.6e-41;
Matches 92; Conservative 11; Mismatches 15; Indels 3; Gaps 1

Qy      1 QVQLKESGPGHVKPSQTLSITCTISGFSLSRYSHVWRQPPGKGLDWLGLMWGGGNTDYN 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 20 QVOLKESGPGVLVAPQSLSITCTVSGFSLTNVAINWVRQPPQGLEWLGIIWTGGTTSYN 79
QY 61 SALKSRLSISKDNKSNQVFLKMNLSLTADTAVYYCARKGEFYGYDGVYWGQGTSLVTVS 120
Db 80 SALKSRLSISKDNKSNQVFLKMNLSLTADTAVYYCARTGTRGYFFD---YWGQGTSLTVS 136
QY 121 S 121
Db 137 S 137

RESULT 24

US-08-652-558-7
: Sequence 7, Application US/08652558
: Patent No. 5861155
: GENERAL INFORMATION:
: APPLICANT: LIN, AUGUSTINE YEE-THARN
: TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
: TITLE OF INVENTION: THEREOF
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WordPerfect 6.1
: CURRENT APPLICATION DATA:
: FILING DATE: JUNE 6, 1996
: APPLICATION NUMBER: US/08/652,558
: PRIORITY DATA:
: FILING DATE: JUNE 6, 1996
: APPLICATION NUMBER: PCT/IB94/00387
: FILING DATE: NOVEMBER 21, 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: YANKWICH, LEON R.
: REGISTRATION NUMBER: 30,237
: REFERENCE/DOCKET NUMBER: 95,497-L
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-345-9100
: TELEFAX: 617-345-9111
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 120 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-652-558-7

Query Match 74.1%; Score 479.5; DB 2; Length 120;
Best Local Similarity 74.0%; Pred. No. 2.2e-41;
Matches 91; Conservative 15; Mismatches 12; Indels 5; Gaps 2;
QY 1 QVOLKESGPGVLVAPQSLSITCTVSGFSLTNVAINWVRQPPQGLEWLGIIWTGGTTSYN 60
Db 1 QVOLKESGPGVLVAPQSLSITCTVSGFSLTNVAINWVRQPPQGLEWLGIIWTGGTTSYN 60
QY 61 SALKSRLSISKDNKSNQVFLKMNLSLTADTAVYYCARK--GEFYGYDGVYWGQGTSLTV 118
Db 61 SALKSRVTMLKDTSKNQVFLRLSSVTAADTAVYYCARDRTATLYAMD---YWGQGSVLT 117
QY 119 VSS 121
Db 118 VSS 120

RESULT 25

US-09-254-189-4

: Sequence 4, Application US/09254189
: Patent No. 6150792
: GENERAL INFORMATION:
: APPLICANT: Lundquist, Tomas
: TITLE OF INVENTION: Sequence Listing
: Patent No. 6150792
: FILE REFERENCE: 3526/00000
: CURRENT APPLICATION NUMBER: US/09/254.189
: CURRENT FILING DATE: 1999-03-02
: EARLIER APPLICATION NUMBER: 00/000.000
: EARLIER FILING DATE: 1998-11-24
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn ver. 2.0
: SEQ ID NO 4
: LENGTH: 120
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:humanized
: OTHER INFORMATION: monoclonal antibody
US-09-254-189-4

Query Match 74.1%; Score 479.5; DB 4; Length 120;
Best Local Similarity 74.0%; Pred. No. 2.2e-41;
Matches 91; Conservative 15; Mismatches 12; Indels 5; Gaps 2;
QY 1 QVOLKESGPGVLVAPQSLSITCTVSGFSLTNVAINWVRQPPQGLEWLGIIWTGGTTSYN 60
Db 1 QVOLKESGPGVLVAPQSLSITCTVSGFSLTNVAINWVRQPPQGLEWLGIIWTGGTTSYN 60
QY 61 SALKSRLSISKDNKSNQVFLKMNLSLTADTAVYYCARK--GEFYGYDGVYWGQGTSLTV 118
Db 61 SALKSRVTMLKDTSKNQVFLRLSSVTAADTAVYYCARDRTATLYAMD---YWGQGSVLT 117
QY 119 VSS 121
Db 118 VSS 120

Search completed: June 23, 2003, 14:04:53
Job time : 8.07979 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:03:36 ; Search time 17.1631 Seconds
(without alignments)
762.858 Million cell updates/sec

Title: US-10-056-052A-20
Perfect score: 647
Sequence: 1 QVQLKESGPGLVKPSQTLSTL.....YYGDFGVYWGQGLTVTVSS 121

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206913 residues

1 number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	647	100.0	121	9	US-10-056-052-20
2	605	93.5	121	9	US-10-056-052-12
3	538.5	83.2	118	9	US-10-056-052-16
4	524.5	81.1	118	9	US-10-056-052-8
5	491.5	76.0	119	12	US-10-140-555-2
6	484.5	74.9	116	9	US-10-194-975-100
7	476	73.6	119	9	US-10-232-187-2
8	474.5	73.3	139	1	US-08-779-784-37
9	472.5	73.0	120	9	US-10-194-975-112
10	471.5	72.9	333	9	US-10-059-261-61
11	468.5	72.4	139	10	US-09-881-823-4
12	468.5	72.4	250	9	US-10-194-975-110
13	467	72.2	119	9	US-10-144-644-15
14	467	72.2	119	9	US-10-144-644-62
15	466.5	72.1	135	1	US-08-779-784-32
16	466.5	72.1	140	10	US-09-007-093-4
17	458	70.8	152	10	US-09-990-205-4
18	457.5	70.7	114	10	US-09-865-483-8
19	454.5	70.2	112	10	US-09-824-286-3

20	447.5	69.2	111	10	US-09-855-271-6	Sequence 6, Appli
21	447	69.1	151	9	US-09-797-941A-2	Sequence 2, Appli
22	445	68.8	474	9	US-10-270-555-3	Sequence 3, Appli
23	443	68.5	249	10	US-09-730-374-3	Sequence 3, Appli
24	436	67.4	144	9	US-09-797-941A-8	Sequence 8, Appli
25	434.5	67.2	116	9	US-10-112-788-7	Sequence 7, Appli
26	434.5	67.2	260	9	US-10-112-788-12	Sequence 12, Appli
27	431	66.6	119	9	US-10-125-687-5	Sequence 5, Appli
28	431	66.6	119	12	US-10-025-687-5	Sequence 5, Appli
29	426	65.8	119	9	US-10-144-644-19	Sequence 19, Appli
30	420.5	65.0	126	9	US-10-067-800-68	Sequence 68, Appli
31	419.5	64.8	135	10	US-09-855-271-28	Sequence 28, Appli
32	413.5	63.9	221	9	US-09-972-656-80	Sequence 80, Appli
33	408	63.1	121	9	US-10-067-800-60	Sequence 60, Appli
34	406.5	62.8	252	9	US-09-880-748-1394	Sequence 1994, Ap
35	406	62.8	251	9	US-09-880-748-990	Sequence 990, App
36	405	62.6	116	10	US-09-881-823-18	Sequence 18, Appli
37	402	62.1	115	9	US-09-144-886-59	Sequence 59, Appli
38	401	62.0	115	9	US-09-144-886-58	Sequence 58, Appli
39	400	61.8	253	9	US-09-880-748-1339	Sequence 1339, Ap
40	399.5	61.7	247	9	US-09-880-748-1651	Sequence 1651, Ap
41	399.5	61.7	252	9	US-09-880-748-1329	Sequence 1329, Ap
42	397.5	61.4	254	9	US-09-880-748-1578	Sequence 1578, Ap
43	397.5	61.4	256	9	US-09-880-748-1745	Sequence 1745, Ap
44	397	61.4	250	9	US-09-880-748-1361	Sequence 1361, Ap
45	396.5	61.3	250	9	US-09-880-748-1548	Sequence 1548, Ap
46	392.5	60.7	139	10	US-09-809-739-9	Sequence 9, Appli
47	392.5	60.7	256	9	US-09-880-748-1607	Sequence 1607, Ap
48	391	60.4	123	9	US-09-995-523-12	Sequence 12, Appli
49	391	60.4	253	9	US-09-880-748-1619	Sequence 1619, Ap
50	390	60.3	119	10	US-09-253-794-9	Sequence 9, Appli
51	390	60.3	119	10	US-09-253-794-11	Sequence 11, Appli
52	390	60.3	119	10	US-09-253-794-57	Sequence 57, Appli
53	389.5	60.2	120	9	US-09-996-288-36	Sequence 36, Appli
54	389.5	60.2	120	9	US-09-996-265-36	Sequence 36, Appli
55	389.5	60.2	450	9	US-09-996-288-220	Sequence 220, App
56	389.5	60.2	450	9	US-09-996-265-220	Sequence 220, App
57	389.5	60.2	450	10	US-09-796-848A-37	Sequence 37, Appli
58	389	60.1	119	10	US-09-253-794-8	Sequence 8, Appli
59	389	60.1	254	9	US-09-880-748-844	Sequence 844, App
60	389	60.1	255	9	US-09-880-748-841	Sequence 841, App
61	388.5	60.0	143	10	US-09-881-823-8	Sequence 8, Appli
62	388	60.0	253	9	US-09-880-748-934	Sequence 934, App
63	387.5	59.9	248	9	US-09-880-748-1360	Sequence 1360, Ap
64	386	59.7	230	9	US-09-972-656-72	Sequence 72, Appli
65	385.5	59.6	249	9	US-09-880-748-1321	Sequence 1321, Ap
66	385.5	59.6	254	9	US-09-880-748-1659	Sequence 1659, Ap
67	385.5	59.6	450	9	US-09-996-288-234	Sequence 232, App
68	385.5	59.6	450	9	US-09-996-288-234	Sequence 234, App
69	385.5	59.6	450	9	US-09-996-265-232	Sequence 232, App
70	385.5	59.6	450	9	US-09-996-265-234	Sequence 234, App
71	385.5	59.6	450	9	US-09-996-265-236	Sequence 236, App
72	385.5	59.6	450	9	US-09-996-265-236	Sequence 236, App
73	384.5	59.4	120	9	US-10-194-975-114	Sequence 114, App
74	384.5	59.4	251	9	US-09-880-748-1538	Sequence 1538, Ap
75	384	59.4	251	9	US-09-880-748-1510	Sequence 1510, Ap
76	383.5	59.3	120	9	US-09-996-288-40	Sequence 40, Appli
77	383.5	59.3	120	9	US-09-996-288-44	Sequence 44, Appli
78	383.5	59.3	120	9	US-09-996-265-40	Sequence 40, Appli
79	383.5	59.3	120	9	US-09-996-265-44	Sequence 44, Appli
80	383.5	59.3	120	10	US-09-796-848A-36	Sequence 36, Appli
81	383.5	59.3	249	9	US-09-880-748-957	Sequence 957, App
82	383.5	59.3	252	9	US-09-880-748-1401	Sequence 1401, Ap
83	383.5	59.3	450	9	US-09-996-288-222	Sequence 222, App
84	383.5	59.3	450	9	US-09-996-288-224	Sequence 224, App
85	383.5	59.3	450	9	US-09-996-265-222	Sequence 222, App
86	383.5	59.3	450	9	US-09-996-265-224	Sequence 224, App
87	383.5	59.3	450	10	US-09-796-848A-39	Sequence 39, Appli
88	383.5	59.3	450	10	US-09-796-848A-41	Sequence 41, Appli
89	383.5	59.3	463	9	US-10-153-382-13	Sequence 13, Appli
90	383	59.2	253	9	US-09-880-748-1602	Sequence 1602, Ap
91	382.5	59.1	120	9	US-09-771-415-18	Sequence 18, Appli
92	382.5	59.1	120	9	US-09-771-415-20	Sequence 20, Appli

93 382.5 59.1 120 9 US-09-771-415-22 Sequence 22, Appl
 94 382.5 59.1 120 9 US-09-771-415-24 Sequence 24, Appl
 95 382.5 59.1 120 9 US-09-771-415-26 Sequence 26, Appl
 96 382.5 59.1 120 9 US-09-996-288-24 Sequence 24, Appl
 97 382.5 59.1 120 9 US-10-011-931-3 Sequence 3, Appl
 98 382.5 59.1 120 9 US-09-996-285-24 Sequence 24, Appl
 99 382.5 59.1 229 10 US-09-974-449-37 Sequence 37, Appl
 100 382.5 59.1 450 9 US-09-996-288-210 Sequence 210, App

ALIGNMENTS

RESULT 1
 US-10-056-052-20
 ; Sequence 20, Application US/10056052
 ; Publication No. US20030099656A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PATTI, Joseph M
 ; APPLICANT: HUTCHINS, Jeff T
 ; APPLICANT: DOMANSKI, Paul
 ; APPLICANT: PATEL, Pratiksha
 ; APPLICANT: HALL, Andrea
 ; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
 ; FILE REFERENCE: P07069US04/BAS
 ; CURRENT FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: 60/308,116
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/298,413
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/274,611
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: 60/264,072
 ; PRIOR FILING DATE: 2001-01-26
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 20
 ; LENGTH: 121
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-10-056-052-20

Query Match 100.0%; Score 647; DB 9; Length 121;
 Best Local Similarity 100.0%; Pred. No. 1.1e-47;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHVVRRQPPGKGLGLEWLGMIWGCGNTDYN 60
 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHVVRRQPPGKGLGLEWLGMIWGCGNTDYN 60
 61 SALKSRLSISKDNSKNOVFLKMNLSLTADTAVYYCARKGEFFYYGVDGFVYWGQGLTLTVTS 120
 61 SALKSRLSISKDNSKNOVFLKMNLSLTADTAVYYCARKGEFFYYGVDGFVYWGQGLTLTVTS 120
 121 S 121
 121 S 121

RESULT 2
 US-10-056-052-12
 ; Sequence 12, Application US/10056052
 ; Publication No. US20030099656A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PATTI, Joseph M
 ; APPLICANT: HUTCHINS, Jeff T
 ; APPLICANT: DOMANSKI, Paul
 ; APPLICANT: PATEL, Pratiksha
 ; APPLICANT: HALL, Andrea
 ; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
 ; FILE REFERENCE: P07069US04/BAS
 ; CURRENT APPLICATION NUMBER: US/10/056,052

; CURRENT FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: 60/308,116
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/298,413
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/274,611
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: 60/264,072
 ; PRIOR FILING DATE: 2001-01-26
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 12
 ; LENGTH: 121
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-10-056-052-12

Query Match 93.5%; Score 605; DB 9; Length 121;
 Best Local Similarity 92.6%; Pred. No. 3.7e-44;
 Matches 112; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHVVRRQPPGKGLGLEWLGMIWGCGNTDYN 60
 1 QVQLKESGPGLVAPVSQSLISITCAISGFSLSRYSVHVVRRQPPGKGLGLEWLGMIWGCGNTDYN 60
 61 SALKSRLSISKDNSKNOVFLKMNLSLTADTAVYYCARKGEFFYYGVDGFVYWGQGLTLTVTS 120
 61 SALKSRLSISKDNSKNOVFLKMNLSLTADTAVYYCARKGEFFYYGVDGFVYWGQGLTLTVTS 120
 121 S 121
 121 A 121

RESULT 3
 US-10-056-052-16
 ; Sequence 16, Application US/10056052
 ; Publication No. US20030099656A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PATTI, Joseph M
 ; APPLICANT: HUTCHINS, Jeff T
 ; APPLICANT: DOMANSKI, Paul
 ; APPLICANT: PATEL, Pratiksha
 ; APPLICANT: HALL, Andrea
 ; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
 ; FILE REFERENCE: P07069US04/BAS
 ; CURRENT APPLICATION NUMBER: US/10/056,052
 ; CURRENT FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: 60/308,116
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/298,413
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/274,611
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: 60/264,072
 ; PRIOR FILING DATE: 2001-01-26
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 16
 ; LENGTH: 118
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-10-056-052-16

Query Match 83.2%; Score 538.5; DB 9; Length 118;
 Best Local Similarity 84.3%; Pred. No. 1.5e-38;
 Matches 102; Conservative 6; Mismatches 10; Indels 3; Gaps 1;
 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHVVRRQPPGKGLGLEWLGMIWGCGNTDYN 60
 1 QVQLKESGPGLVAPVSQSLISITCTISGFSLSRYSVHVVRRQPPGKGLGLEWLGMIWGCGNTDYN 60
 61 SALKSRLSISKDNSKNOVFLKMNLSLTADTAVYYCARKGEFFYYGVDGFVYWGQGLTLTVTS 120


```
US-10-059-261-61
; Sequence 61, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS
; APPLICANT: BRIAND, JEAN-PAUL
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
; TITLE OF INVENTION: SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOTIC
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
; TITLE OF INVENTION: (PTPC)
; FILE REFERENCE: 03495.0216
; CURRENT APPLICATION NUMBER: US/10/059,261
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/265,594
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector pACgp67-ScFv461
; OTHER INFORMATION: peptide sequence
US-10-059-261-61

Query Match 72.9%; Score 471.5; DB 9; Length 333;
Best Local Similarity 74.0%; Pred. No. 1.7e-32;
Matches 91; Conservative 12; Mismatches 15; Indels 5; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRPPGKGLWLGIMWGNDYDYN 60
DB 86 QVQLKESGPGLVAPSSQSLTCTVSGFSLTNYDINWVRPPGKGLWLGIMWGDSNYH 145

QY 61 SALKSRLSISKDKNQVFLKMNLSLTAAADTAVYYCARKGEF--YYGVDGFFVWGQGLT 118
DB 146 SALISRLSISKDKNQVFLKMNLSLTAAADTAVYYCARKGEF--YYGVDGFFVWGQGLT 202

QY 119 VSS 121
DB 203 VSS 205

RESULT 11
US-09-881-823-4
; Sequence 4, Application US/09881823
; Patent No. US20020068066A1
; GENERAL INFORMATION:
; APPLICANT: SHI, WENYUAN
; APPLICANT: ANDERSON, MAXWELL
; APPLICANT: MORRISON, SHERIE
; APPLICANT: TRINH, RYAN
; APPLICANT: WIMS, LETITIA
; APPLICANT: CHEN, LI
; TITLE OF INVENTION: Method for the Treatment and Prevention of Dental Caries
; FILE REFERENCE: 22851-032
; CURRENT APPLICATION NUMBER: US/09/881,823
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 07/378,577
; PRIOR FILING DATE: 1999-08-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Murine
US-09-881-823-4

Query Match 72.4%; Score 468.5; DB 10; Length 139;
Best Local Similarity 72.7%; Pred. No. 1.4e-32;
Matches 88; Conservative 14; Mismatches 16; Indels 3; Gaps 1;
```

```
QY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRPPGKGLWLGIMWGNDYDYN 60
DB 20 QVQLKESGPGLVAPSSQSLTCTVSGFSLTNYDINWVRPPGKGLWLGIMWGDSNYH 79

QY 61 SALKSRLSISKDKNQVFLKMNLSLTAAADTAVYYCARKGEFYYGVDGFFVWGQGLT 120
DB 80 SALISRLSISKDKNQVFLKMNLSLTAAADTAVYYCARKGEFYYGVDGFFVWGQGLT 136

QY 121 S 121
DB 137 S 137

RESULT 12
US-10-194-975-110
; Sequence 110, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 110
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: humanized chimeric D1.3 antibody
US-10-194-975-110

Query Match 72.4%; Score 468.5; DB 9; Length 250;
Best Local Similarity 73.6%; Pred. No. 2.3e-32;
Matches 89; Conservative 16; Mismatches 11; Indels 5; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRPPGKGLWLGIMWGNDYDYN 60
DB 24 QVQLKESGPGLVKPSQTLSTCTVSGGSGVGNWIRPPGKGLWLGIMWGNDYDYN 83

QY 61 SALKSRLSISKDKNQVFLKMNLSLTAAADTAVYYCARKGEFYYGVDGFFVWGQGLT 120
DB 84 SSKSRVTISVDISKQNFSLKLSSTAAADTAVYYCARERD--YRLD---YWGQGLT 138

QY 121 S 121
DB 139 S 139

RESULT 13
US-10-144-644-15
; Sequence 15, Application US/10144644
; Publication No. US20030059429A1
; GENERAL INFORMATION:
; APPLICANT: Ames, Robert S.
; APPLICANT: Appelbaum, Edward R.
; APPLICANT: Chaiken, Irwin M.
; APPLICANT: Cook, Richard M.
; APPLICANT: Gross, Mitchell S.
; APPLICANT: Holmes, Stephen D.
; APPLICANT: McMillan, Lynette J.
; APPLICANT: Theisen, Timothy W.
; TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
; TITLE OF INVENTION: Treatment of IL5 Mediated Disorders
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corp
; STREET: P.O. Box 1539-UW2220
; CITY: King of Prussia
; STATE: Pennsylvania
```

```

1 COUNTRY: USA
2
3 ZIP: 19406-0939
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: PatentIn Release #1.0, Version #1.30
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/10/144,644
13 FILING DATE: 2002-08-08
14
15 CLASSIFICATION:
16
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: US 08/470110
19 FILING DATE: 06-JUN-1995
20
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 08/467420
23 FILING DATE: 06-JUN-1995
24
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: US 08/363131
27 FILING DATE: 23-DEC-1994
28
29 ATTORNEY/AGENT INFORMATION: INFORMATION:
30 NAME: Sutton, Jeffrey A.
31
32 REGISTRATION NUMBER: 34,028
33
34 REFERENCE/DOCKET NUMBER: P50282-2
35
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: 610-270-5024
38
39 TELEFAX: 610-270-5090
40
41 INFORMATION FOR SEQ ID NO: 15:
42 SEQUENCE CHARACTERISTICS:
43 LENGTH: 119 amino acids
44 TYPE: amino acid
45 STRANDEDNESS: single
46 TOPOLOGY: linear
47
48 MOLECULE TYPE: protein
49
50 US-10-144-644-15

```

	Query Match	72.2%	Score 467;	DB 9;	Length 119;
	Best Local Similarity	72.4%;	Pred. No. 1.6e-32;		
Matches	92; Conservative	8;	Mismatches 13;	Indels 14;	Gaps 2;
QY	1 QVQLKESGPGCLVKPSSOTLSITCTISGFSLSRYSVHWVRPPGKGLEWIGMTWGSGNTDYN	60			
Db	1 QVQLKESGPGCLVAPSOSLSTCTVSGFSLTSYSVHWVRPPGKGLEWLGVIAWSGGTDYN	60			
QY	61 SALKSRLSISKDNKSQNQVFLKMNSLTAAATAVYYCARKGEFYGYD-----GFVYWGQG	114			
Db	61 SALMSRLSISKDNKSQVFLKLNSLODTDTAMYYCAR-----DPSSLLRLDYWGQG	112			
	115 TLTVTSS 121				
	:				
	113 TLTVTSS 119				
	:				

RESULT 14
US-10-144-644-62
; Sequence 62, Application US/10144644
; Publication No. US20030059429A1
; GENERAL INFORMATION:
; APPLICANT: Ames, Robert S.
; APPLICANT: Appelbaum, Edward R.
; APPLICANT: Chaiken, Irwin M.
; APPLICANT: Cook, Richard M.
; APPLICANT: Gross, Mitchell S.
; APPLICANT: Holmes, Stephen D.
; APPLICANT: McMillan, Lynette J.
; APPLICANT: Theisen, Timothy W.
; TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
; TITLE OF INVENTION: Treatment of IL5 Mediated Disorders
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corp
; STREET: P.O. Box 1539-UW2220
; CITY: King of Prussia

```

; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/144,644
; FILING DATE: 2002-08-08
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/470110
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/467420
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363131
; FILING DATE: 23-DEC-1994
; ATTORNEY/AGENT INFORMATION: INFORMATION:
; NAME: Sutton, Jeffrey A.
; REGISTRATION NUMBER: 34,028
; REFERENCE/DOCKET NUMBER: P50282-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5024
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-10-144-644-62

Query Match 72.2% Score 467; DB 9; Length 119;
Best Local Similarity 71.7%; Pred. No. 1.6e-32;
Matches 91; Conservative 11; Mismatches 11; Indels 14; Gaps 2%

QY 1 QVQLKESGPGLVKPSQTLSITCTISGFSLSRYSVHVVRRQPPGKLEWLGMIWGGNTDYN 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 QVQLQESGPGLVKPSQTLSITCTISGFSLSRYSVHVVRRQPPGKLEWLGVIWASGGTDYN 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 SALKSRLSISKDNSKNQVFLKMSLTAADTAVVYVCARKGEYYGYD-----GFVYWCQ 114
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 SALSMSRLSILKDNSKNQVSLRLSVTAADTAVVYCAR-----DPPSSLLRLLDYWGQ 112
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 115 TLTVTVSS 121
   |||||
Db 113 TTVTVSS 119
   |||||

RESULT 15
US-08-779-784-32
; Sequence 32, Application US/08779784
; Patent No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiko
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; NUMBER OF INVENTION: REMELINATION USING MONOCLONAL AUTOANTIBODIES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/779,784
APPLICATION NUMBER: US/08/779,784
FILING DATE: 07-JAN-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/692,084
FILING DATE: 08-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,520
FILING DATE: 29-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-779-784-32

Query Match 72.1%; Score 466.5; DB 1; Length 135;
Best Local Similarity 75.2%; Pred. No. 2e-32;
Matches 91; Conservative 10; Mismatches 15; Indels 5; Gaps 2;
Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGMIWGSGTDYN 60
Db 20 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGMIWGSGTDYN 79
Qy 61 SALKSRISKDNKSNQVFLKMSLTAAADTAIVYCARKEGFYGYDGFVYWGQGLTVTS 120
Db 80 AAFISRLSISKDNKSNQVFLKMSLTAAADTAIVYCARKEGFYGYDGFVYWGQGLTVTS 134
121 S 121
135 S 135

RESULT 16
US-09-007-093-4
Sequence 4, Application US/09007093
Patent No. US2002025315A1
GENERAL INFORMATION:
APPLICANT: Anand, Naveen N
APPLICANT: Barber, Brian H
APPLICANT: Cates, George A
APPLICANT: Caterini, Judith E
APPLICANT: Klein, Michel H
TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF
TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/007,093
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/483,576
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-765
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-007-093-4

Query Match 72.1%; Score 466.5; DB 10; Length 140;
Best Local Similarity 74.0%; Pred. No. 2e-32;
Matches 91; Conservative 12; Mismatches 15; Indels 5; Gaps 3;
Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGMIWGSGTDYN 60
Db 21 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGMIWGSGTDYN 80
Qy 61 SALKSRISKDNKSNQVFLKMSLTAAADTAIVYCARKEGFYGYDGFVYWGQGLTVT 118
Db 81 SALKSRISKDNKSNQVFLKMSLTAAADTAIVYCARKEGFYGYDGFVYWGQGLTVT 137
Qy 119 VSS 121
Db 138 ASS 140

RESULT 17
US-09-990-205-4
Sequence 4, Application US/09990205
Patent No. US20020150572A1
GENERAL INFORMATION:
APPLICANT: FOON, Kenneth A.
APPLICANT: CHATTERJEE, Malaya
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF PSORIASIS
FILE REFERENCE: 304142000501
CURRENT APPLICATION NUMBER: US/09/990,205
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: U.S. 09/192,838
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: U.S. 60/065,774
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 152
TYPE: PRT
ORGANISM: Mus Musculus
US-09-990-205-4
Query Match 70.8%; Score 458; DB 10; Length 152;
Best Local Similarity 72.7%; Pred. No. 1.1e-31;
Matches 88; Conservative 10; Mismatches 19; Indels 4; Gaps 1;
Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGMIWGSGTDYN 60
Db 20 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRQPPGKGLWLGMIWGSGTDYN 79

GENERAL INFORMATION:
APPLICANT: ELLIS, JONATHAN H.
LEWIS, ALAN P.
TITLE OF INVENTION: HUMANISED ANTIBODIES TO CD38
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 Mb disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/797,941A
FILING DATE: 05-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/849,410
FILING DATE: 30-MAY-1997
APPLICATION NUMBER: PCT/GB95/02777
FILING DATE: 28-NOV-1995
APPLICATION NUMBER: GB 9424449.8
FILING DATE: 02-DEC-1994
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-797-941A-2

Query Match 69.1%; Score 447; DB 9; Length 151;
Best Local Similarity 70.7%; Pred. No. 9.5e-31;
Matches 87; Conservative 12; Mismatches 18; Indels 6; Gaps 2;
Qy 1 QVQLKESGPGLVKPSQTLTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
Db 22 QVQLKESGPGLVKPSQTLTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 81
Qy 61 SALKSRLSISKDKNKQVFLKMNLSLTADTAIVYVCARKGEFYGYDGFVY--WGQGLT 118
Db 82 AAFMSRLNITKDNKQVFLKMNLSLTADTAIVYVCARKGEFYGYDGFVY--WGQGLT 137
Qy 119 VSS 121
Db 138 VSS 140

RESULT 22
US-10-270-555-3
Sequence 3, Application US/10270555
Publication No. US20030092068A1
GENERAL INFORMATION:
APPLICANT: Asahi Kasei Corporation
TITLE OF INVENTION: Agents for adsorption and cross-linkage for adenovirus
FILE REFERENCE: PH-16480S
CURRENT APPLICATION NUMBER: US/10/270,555
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: JP 2001-317766
PRIOR FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 13
SEQ ID NO 3
LENGTH: 474
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: a fusion protein of human CAR and a single chain Fv derived from

; OTHER INFORMATION: monoclonal antibody against human CD34 with an artificial link
US-10-270-555-3
Query Match 68.8%; Score 445; DB 9; Length 474;
Best Local Similarity 71.1%; Pred. No. 4e-30;
Matches 86; Conservative 11; Mismatches 20; Indels 4; Gaps 2;
Qy 1 QVQLKESGPGLVKPSQTLTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
Db 230 QVQLKESGPGLVKPSQTLTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 289
Qy 61 SALKSRLSISKDKNKQVFLKMNLSLTADTAIVYVCARKGEFYGYDGFVYWGQGLTVTS 120
Db 290 AAFISRLSISRDISKQVFFKMNLSQVDDTAIYYCARNR--YESY--FDYWGQGLTVTS 345
Qy 121 S 121
Db 346 S 346
RESULT 23
US-09-730-374-3
Sequence 3, Application US/09730374
Patent No. US20010031126A1
GENERAL INFORMATION:
APPLICANT: Donovan, Kathleen A.
APPLICANT: Lust, John A.
TITLE OF INVENTION: USE OF GENETICALLY ENGINEERED ANTIBODIES
TO CD38 TO TREAT MULTIPLE MYELOMA
FILE REFERENCE: 150.188US2
CURRENT APPLICATION NUMBER: US/09/730,374
CURRENT FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: PCT/US99/12512
PRIOR FILING DATE: 1999-06-04
PRIOR APPLICATION NUMBER: 60/088,277
PRIOR FILING DATE: 1998-08-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 249
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: A polypeptide encoded by an open reading frame of
US-09-730-374-3
Query Match 68.5%; Score 443; DB 10; Length 249;
Best Local Similarity 71.9%; Pred. No. 3.3e-30;
Matches 87; Conservative 11; Mismatches 21; Indels 2; Gaps 2;
Qy 1 QVQLKESGPGLVKPSQTLTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
Db 7 KVLQESGPGSLVQPSQRLSITCTVSGFSLISYGVHWVRQSPGKGLWLGVIWRGGSDYN 66
Qy 61 SALKSRLSISKDKNKQVFLKMNLSLTADTAIVYVCARKGEFYGYDGFVYWGQGLTVTS 120
Db 67 AAFMSRLSITKDNKSKQVFFKMNLSQVDDTAIYYFCA-KTLITTCY-ANDYWGQGLTVTS 124
Qy 121 S 121
Db 125 S 125
RESULT 24
US-09-797-941A-8
Sequence 8, Application US/09797941A
Patent No. US20020164788A1
GENERAL INFORMATION:
APPLICANT: ELLIS, JONATHAN H.
LEWIS, ALAN P.
TITLE OF INVENTION: HUMANISED ANTIBODIES TO CD38
NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 Mb disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/797,941A
FILING DATE: 05-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/849,410
FILING DATE: 30-MAY-1997
APPLICATION NUMBER: PCT/GB95/02777
FILING DATE: 28-NOV-1995
APPLICATION NUMBER: GB 9424449.8
FILING DATE: 02-DEC-1994
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 144 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-797-941A-8

Query Match 67.4%; Score 436; DB 9; Length 144;
Best Local Similarity 68.6%; Pred. No. 7.7e-30;
Matches 81; Conservative 19; Mismatches 12; Indels 6; Gaps 2;
Qy 1 QVQLKESGPGLVKPSQSLTCTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGSGNTDYN 60
Db 31 QVQLKESGPGLVKPSQSLTCTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGSGNTDYN 90
Qy 61 SALKSRSLISKDNKSNQVFLKMNLSLTADTAADTAAYVCARKGEFYGYDGFVY--WGQGTLL 116
Db 91 AAFMSRLNITKDNKSNQVFLKMNLSLTADTAADTAAYVCARKS-----MITTGFVMDSNQGGSL 144

RESULT 25
US-10-112-788-7
Sequence 7, Application US/10112788
Application No. US2003007766A1
GENERAL INFORMATION:
APPLICANT: DENARDO, SALLY
APPLICANT: WINTHROP, MICHELLE
APPLICANT: DENARDO, GERALD
TITLE OF INVENTION: ANTI-MUC-1 SINGLE CHAIN ANTIBODIES FOR TUMOR TARGETING
FILE REFERENCE: 309T-000210US
CURRENT APPLICATION NUMBER: US/10/112.788
CURRENT FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 60/280,721
PRIOR FILING DATE: 2001-04-30
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 116
TYPE: PRT
ORGANISM: Mus musculus
US-10-112-788-7

Query Match 67.2%; Score 434.5; DB 9; Length 116;
Best Local Similarity 66.1%; Pred. No. 8.4e-30;
Matches 80; Conservative 16; Mismatches 20; Indels 5; Gaps 1;
Qy 1 QVQLKESGPGLVKPSQSLTCTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGSGNTDYN 60
Db 91 AAFMSRLNITKDNKSNQVFLKMNLSLTADTAADTAAYVCARKS-----MITTGFVMDSNQGGSL 144

Db 1 QVQLKESGPGLVKPSQSLTCTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGSGNTDYN 60
Qy 61 SALKSRSLISKDNKSNQVFLKMNLSLTADTAADTAAYVCARKGEFYGYDGFVYWGQGTLLTVTS 120
Db 61 PAFISRLNINKDNKSNQVFLKMNLSLTADTAADTAAYVCARKGEFYGYDGFVYWGQGTLLTVTS 115
Qy 121 S 121
Db 116 S 116

Search completed: June 23, 2003, 14:19:14
Job time : 18.1631 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 19.5231 Seconds
(without alignments)
825,861 Million cell updates/sec.

Title: US-10-056-052A-20

Perfect score: 647

Sequence: 1 QVQLKESGFLVKPSQTLSI.....YGYDGFVYWGQGLTVTVSS 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

al number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_101002.*

```

1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523.5	80.9	118	22	AA1981970
2	523.5	80.9	581	22	AA1981972
3	506	78.2	119	22	AA198675
4	506	78.2	138	22	AA198656
5	502	77.6	113	22	AA1982025
6	496	76.7	121	21	AA1980020
7	496	76.7	476	22	AA1984243
8	494.5	76.4	137	18	AA1980273
9	494	76.4	183	12	AA1981526
10	494	76.4	183	14	AA1982128
11	468	72.4	116	17	AA1981490
12	468	72.4	116	17	AA1981490
13	468	72.4	116	17	AA1981490
14	468	72.4	116	17	AA1981490
15	468	72.4	116	17	AA1981490
16	468	72.4	116	17	AA1981490
17	468	72.4	116	17	AA1981490
18	468	72.4	116	17	AA1981490
19	468	72.4	116	17	AA1981490
20	468	72.4	116	17	AA1981490
21	468	72.4	116	17	AA1981490
22	468	72.4	116	17	AA1981490
23	468	72.4	116	17	AA1981490
24	468	72.4	116	17	AA1981490
25	468	72.4	116	17	AA1981490
26	468	72.4	116	17	AA1981490
27	468	72.4	116	17	AA1981490
28	468	72.4	116	17	AA1981490
29	468	72.4	116	17	AA1981490
30	468	72.4	116	17	AA1981490
31	468	72.4	116	17	AA1981490
32	468	72.4	116	17	AA1981490
33	468	72.4	116	17	AA1981490
34	468	72.4	116	17	AA1981490
35	468	72.4	116	17	AA1981490
36	468	72.4	116	17	AA1981490
37	468	72.4	116	17	AA1981490
38	468	72.4	116	17	AA1981490
39	468	72.4	116	17	AA1981490
40	468	72.4	116	17	AA1981490
41	468	72.4	116	17	AA1981490
42	468	72.4	116	17	AA1981490
43	468	72.4	116	17	AA1981490
44	468	72.4	116	17	AA1981490
45	468	72.4	116	17	AA1981490
46	468	72.4	116	17	AA1981490
47	468	72.4	116	17	AA1981490
48	468	72.4	116	17	AA1981490
49	468	72.4	116	17	AA1981490
50	468	72.4	116	17	AA1981490
51	468	72.4	116	17	AA1981490
52	468	72.4	116	17	AA1981490
53	468	72.4	116	17	AA1981490
54	468	72.4	116	17	AA1981490
55	468	72.4	116	17	AA1981490
56	468	72.4	116	17	AA1981490
57	468	72.4	116	17	AA1981490
58	468	72.4	116	17	AA1981490
59	468	72.4	116	17	AA1981490
60	468	72.4	116	17	AA1981490
61	468	72.4	116	17	AA1981490
62	468	72.4	116	17	AA1981490
63	468	72.4	116	17	AA1981490
64	468	72.4	116	17	AA1981490
65	468	72.4	116	17	AA1981490
66	468	72.4	116	17	AA1981490
67	468	72.4	116	17	AA1981490
68	468	72.4	116	17	AA1981490
69	468	72.4	116	17	AA1981490
70	468	72.4	116	17	AA1981490
71	468	72.4	116	17	AA1981490
72	468	72.4	116	17	AA1981490
73	468	72.4	116	17	AA1981490
74	468	72.4	116	17	AA1981490
75	468	72.4	116	17	AA1981490
76	468	72.4	116	17	AA1981490
77	468	72.4	116	17	AA1981490
78	468	72.4	116	17	AA1981490
79	468	72.4	116	17	AA1981490
80	468	72.4	116	17	AA1981490
81	468	72.4	116	17	AA1981490
82	468	72.4	116	17	AA1981490
83	468	72.4	116	17	AA1981490

Mouse germ-line co
Heavy chain variab
Anti-NKG2D hybrid
Murine consensus 1
Humanised anti-CD4
Ganglioside GD2 sp
Ganglioside GD2 sp
Sequence of the he
161-107/113-60/hl-
Peptide correspond
Antibody scfv8 hea
TM27 antibody VH c
ECOR1-HindIII inse
VLLys-linker-VHlys
VHDI.3-HuCH1 regio
18-2-3-TRY59. A
Single chain polyp
18-2-3TRY59 singl
Single chain bindi
Two linked VHLYs p
SCFV sequences enc
SCFV anti-lysozyme
Sequence translate
Murine 5D12 monocl
Heavy chain variab
Heavy chain variab
Amino acid sequenc
VH NQ2/12.4-VK NQ1
VH NQ2/12.4-VK NQ1
Sequence of the si
PUR4125 VL-Lys-syn
Anti-DNA antibody
Single chain antib
SCFV(225)-ETA fusi
scfv2(fRP5/225)-ET
scfv2(fRP5/225)-ET
scfv2(fRP5/225)-ET
18-2-3-TRY202.
166-111/112-65/hl-
Fusion protein enc
Heavy chain of MUM
TM27 antibody VH c
pectate lyase lead
Murine anti-Lewis
Single chain polyp
Single chain bindi
Murine VH group 2
Humanised anti-CTL
18-2-3TRY202' sin
Human anti-CD40 mo
MAB 10.1 heavy cha
Truncated SCFV ant
Humanised ID10 ant
Humanised ID10 ant
Humanised ID10 ant
Anti-NKG2D hybrid
Anti-NKG2D hybrid
6E5A74-7 bispecif
Human p53 tetramer
P5-9 single chain
M79scfv-interleuki
M79scfv-interleuki
VH domain of antib
Monoclonal antibod
Monoclonal antibod
NEMM humanised 2B6
Mouse anti-human I
TM27 antibody VH c
Murine anti-human
Murine anti-human
Murine 4H104 mab
TM27 antibody VH c
NEMM humanised hea


```

XX 23-MAR-1998 (first entry)
XX
XX Humanized 5D12 monoclonal antibody heavy chain variable region.
XX
XX Human: B7 antigen; CD40; monoclonal antibody; B cell; growth;
XX differentiation; allergy; autoimmune disease.
XX
XX Chimeric - Mus sp.
XX Chimeric - Homo sapiens.
XX
XX WO9731025-A1.
XX
XX 28-AUG-1997.
XX
XX 21-FEB-1997; 97WO-US02858.
XX
XX 23-FEB-1996; 96US-0606293.
XX
XX (CHIR ) CHIRON CORP.
XX
XX De Boer M;
XX
XX WPI: 1997-435094/40.
XX N-PSDB; AAT79192.
XX
XX Humanised monoclonal antibody binding CD40 antigen on B cells - to
XX prevent their growth and differentiation, useful to treat or prevent
XX antibody mediated allergies and autoimmune diseases
XX
XX Example 8; Fig 12; 64pp; English.
XX
XX A novel humanised monoclonal antibody (MAB) has been developed which:
XX (a) binds a human CD40 antigen on the surface of a B cell, to prevent
XX its growth and differentiation; and (b) has an effective number of
XX exposed amino acids in its framework regions that are consistent with
XX amino acid residues found in the corresponding framework regions of a
XX human Ab to provide a reduced immunogenicity in humans. The present
XX sequence represents humanized 5D12 monoclonal antibody heavy chain
XX variable region, used in an example of the present invention. The MAB
XX can be used to prevent or treat antibody mediated diseases, particularly
XX immunoglobulin E (IgE) mediated allergies, systemic lupus erythematosus,
XX primary biliary cirrhosis, idiopathic thrombocytopenia purpura and
XX rheumatoid arthritis. The MAB inhibits the normal B cell response to
XX CD40 ligands at relatively low concentrations and is unlikely to cause
XX any immune response in humans.
XX
XX SQ Sequence 113 AA;
XX
XX Query Match 77.6%; Score 502; DB 18; Length 113;
XX Best Local Similarity 81.0%; Pred. No. 1.4e-39;
XX Matches 98; Conservative 7; Mismatches 8; Indels 8; Gaps 2;
XX
XX 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGIMGGNTDYN 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 1 QVQLVESGPGLVKPSQSLSITCTVSGFSLSRYSVHWVRQPPGKGLWLGIMGGSTDYN 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 61 SALKSRSLTSKDNKSNQVFLKMNLSLTAAATVYYCYCARK-GEFYGYGQGLTVTVS 120
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 61 SALKSRSLTSKDTSKNOVFLKMNLSLRAEDTAMYYCVRT-----DG-DYWGQGLTVTVS 112
XX 121 S 121
XX 113 S 113
XX
XX RESULT 6
XX AAB10020
XX ID AAB10020 standard; Protein; 121 AA.
XX
XX AAB10020;
XX
XX 01-NOV-2000 (first entry)
XX
XX DT

```

```

XX H. pylori HSP60-binding antibody heavy chain protein.
XX
XX Acid-resistant microorganism; detection: faecal; intestine; infection;
XX monoclonal antibody; heat shock protein; HSP60; heavy chain.
XX
XX Unidentified.
XX
XX WO200026671-A1.
XX
XX 11-MAY-2000.
XX
XX 29-OCT-1999; 99WO-EP08212.
XX
XX 29-OCT-1998; 98EP-0120517.
XX 06-NOV-1998; 98EP-0120687.
XX
XX (CONN-) CONNEX GMBH.
XX
XX Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;
XX Ringels A;
XX
XX WPI: 2000-365747/31.
XX N-PSDB; AAA40200.
XX
XX Detecting infection by acid-fast microbes for diagnosis of Helicobacter
XX pylori, comprises reacting a faecal sample with two binding reagents for
XX antigens that survive intestinal passage -
XX
XX Disclosure; Fig 2; 84pp; German.
XX
XX This invention describes a novel method for the detection of a mammalian
XX infection by an acid-resistant microorganism (A) by treating a faecal
XX sample with at least two different monoclonal antibodies (WAB) (or their
XX fragments or derivatives) or aptamers (collectively (I)) and detecting
XX formation of a complex (C) between (I) and the corresponding antigen of
XX (A). The first and second (I) bind to epitopes of different antigens
XX (Ag). These epitopes are present, after passage through the intestines,
XX in at least some mammals, and have either: (i) their native structure;
XX or (ii) a structure against which an antibody is produced by an animal
XX infected or immunized with (A), or its extract, lysate, derived protein
XX or fragment, or with a synthetic peptide. Practically all mammals display
XX at least one of the specified epitopes. The method is used to detect
XX infection by acid-fast bacteria, particularly of the genera Helicobacter,
XX Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,
XX M. tuberculosis, C. jejuni and C. pylori. (I) may also be used
XX therapeutically. The method is direct and non-invasive, and provides an
XX inexpensive and easily standardizable diagnosis, despite possible
XX degradation of antigens during passage through the intestines. This
XX sequence represents the H. pylori heat shock protein, HSP60-binding
XX antibody (DMS ACC2356) heavy chain which is used to illustrate the method
XX of the invention.
XX
XX SQ Sequence 121 AA;
XX
XX Query Match 76.7%; Score 496; DB 21; Length 121;
XX Best Local Similarity 79.7%; Pred. No. 5.6e-39;
XX Matches 98; Conservative 8; Mismatches 13; Indels 4; Gaps 3;
XX
XX 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGIMGGNTDY 59
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 1 EVQLLEESGPGLVAPQSQSLSTCTVSGFSLSRYSVHWVRQPPGKGLWLGIMGGSTDY 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 60 NSALKSRSLTSISKDNKSNQVFLKMNLSLTAAATVYYCYCARK-GEFYGYGQGLTVLT 118
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 61 NSGLKSRSLTSISNDNSKSKVFLKMNLSLTQDQDTAIYYCARNMGGRYPDY--FDYWGQGLT 118
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 119 VSS 121
XX 119 VSS 121
XX
XX RESULT 7

```

```

AAB49243
ID AAB49243 standard; protein; 476 AA.
XX
XX AAB49243;
AC
XX
XX
DT 15-MAR-2001 (first entry)
XX
DE Chimeric 4H6 anti-DR4 antibody heavy chain protein.
XX
XX Anti-Death receptor 4; DR4; antibody; apoptosis; cancer; arthritis;
KW autoimmune.
KW
XX Homo sapiens.
OS Synthetic.
XX
XX WO200073349-A1.
XX
XX 07-DEC-2000.
PD
XX
XX 25-MAY-2000; 2000WO-US14599.
XX
XX 28-MAY-1999; 99US-0322875.
XX
XX (GETH ) GENENTECH INC.
PA
XX Ashkenazi AJ, Chuntharapai A, Dodge KH, Kim KJ;
XX
XX WPI; 2001-041145/05.
XX
XX Novel anti-death receptor 4 antibodies useful for treating cancer and
PT immune related disorders such as rheumatoid arthritis, sjogren's
PT syndrome, Grave's disease and diabetes mellitus -
XX
XX Claim 16; Fig 18; 136pp; English.
XX
XX The present invention relates to an anti-Death receptor 4 (DR4)
CC antibody. The antibodies of the invention are useful for inducing
CC apoptosis in mammalian cancer cells such as colon cancer cells and
CC for treating an immune-related disease in a mammal such as arthritis
CC and autoimmune disease.
XX
XX Sequence 476 AA;
XX
Query Match 76.7%; Score 496; DB 22; Length 476;
Best Local Similarity 76.2%; Pred. No. 2:6e-38;
Matches 96; Conservative 10; Mismatches 14; Indels 6; Gaps 2;
O 2 VOLKESGGLYKPSQTLSTICTISGFSLSRYSVHVRPPGKGLWLMGCGGNTDYS 61
21 VOLKESGGLVAPSOGLSTICTVSGFSLTSYGVHVRPPGKGLWLVAVGSTNYS 80
OY 62 ALKSLRSLISKDKNQVFLKNSLTAADTAVVYCARKGEF-YYG-----YDGFVYWGQT 115
II IIIIIIIIIIIIIIIIIIIII IIIIIIIIIIIII IIIIIIIIIIIII
Db 81 ALMSRLSLISKDKNQVFLKNSLTDOTDAMYCARCEGEFDYGGSSLLSYHSMNFWGGQT 140
OY 116 LVTVSS 121
IIIIII
Db 141 SVTVSS 146
IIIIII

RESULT 8
ID AAW30273
XX AAW30273 standard; Protein; 137 AA.
XX
XX AAW30273;
AC
XX
XX 07-JUL-1998 (first entry)
DT
XX Heavy chain of HuM4TS.22.
DE
XX HuM4TS.22; antibody; platelet derived growth factor beta receptor; PDGF-R beta;
KW inhibition; intimal hyperplasia; vasculature; restenosis; angioplasty;
KW heavy chain.

```

Query Match 76.4%; Score 494.5; DB 18; Length 137;

Best Local Similarity 76.9%; Pred. No. 8.9e-39;
Matches 93; Conservative 14; Mismatches 11; Indels 3; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSILSRYSVHVVRPPCKGLEWLGMIWGCGNTDYN 60
DB 20 QVQLKESGPGLVKPSQTLISITCTISGFSILSRYSVHVVRPPCKGLEWLGMIWGCGNTDYN 79
QY 61 SALKSRLSISKDNSKNQVFLKMNLSLTAADTAVYYCARKGEFYGYDG--FYWGQGLTVTS 120
DB 80 SALKSRLSISKDTSKNQVSLKSLSSVTAADTAVYYCARTGRCYFED---YWGQGLTVTS 136
QY 121 S 121
DB 137 S 137

RESULT 9
AAR15326
ID AAR15326 standard; Protein: 183 AA.

AAR15326;

DT 16-MAR-1992 (first entry)

XX IL-2 chimeric antibody heavy chain.
XX Interleukin-2; immunosuppressant.
XX Chimeric Homo sapiens.
OS Chimeric Mus musculus.

XX Key Location/Qualifiers
FT Peptide 1..19
FT /note= "signal peptide"
FT Peptide 20..116
FT /note= "V-region"
FT Peptide 117..123
FT /note= "D-region"
FT Peptide 124..138
FT /note= "J3-region"
FT Peptide 139..183
FT /note= "C-region"

XX EP460674-A.

PN 11-DEC-1991.

XX 06-JUN-1991; 91EP-0109303.

XX 08-JUN-1990; 90DE-4018442.

XX (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Weidle U, Kaluza B, Knapp W;

XX WPI; 1991-363168/50.

DR N-PSDB; AAQ15118.

XX New recombinant DNA encoding chimeric antibody - with human
PT constant and non-human variable regions, esp. directed against
PT interleukin 2 receptor

XX Disclosure; Page 15; 24pp; German.

XX The sequence is that of the interleukin-2 (IL-2) chimeric antibody
CC heavy chain, coded by clone 215. Antibodies against IL-2 receptors
CC are useful as immunosuppressants, and are much less immunogenic than
CC mouse or rat antibodies. See also AAR15321-R15325.

XX Sequence 183 AA;

Query Match 76.4%; Score 494; DB 12; Length 183;
Best Local Similarity 77.2%; Pred. No. 1.4e-38;

Matches 95; Conservative 10; Mismatches 12; Indels 6; Gaps 2;
QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSILSRYSVHVVRPPCKGLEWLGMIWGCGNTDYN 60
DB 20 QVQLKESGPGLVKPSQTLISITCTISGFSILSRYSVHVVRPPCKGLEWLGMIWGCGNTDYN 79
QY 61 SALKSRLSISKDNSKNQVFLKMNLSLTAADTAVYYCARKGEFYGYDG--FYWGQGLTVT 118
DB 80 SALKSRLSISKDTSKNQVSLKSLSSVTAADTAVYYCARTGRCYFED---YWGQGLTVT 135
QY 119 VSS 121
DB 136 VSA 138

RESULT 10
AAR32128
ID AAR32128 standard; Protein: 183 AA.

XX AAR32128;

XX 02-JUN-1993 (first entry)

XX Anti-IL2R-alpha antibody M-215 heavy chain variable region.

XX immunosuppression; tissue transplantation; graft; L chain; V region;
KW T-helper cell inhibition; transplant rejection; MAB;
KW interleukin-2 receptor.

XX Key Location/Qualifiers
FT Peptide 1..19
FT /label= signal
FT Region 20..116
FT /label= Variable-region
FT Region 117..123
FT /label= D-region
FT Region 124..138
FT /label= J3-region
FT Region 139..183
FT /label= start_of_constant_region

XX DE4143214-A.

PN 28-JAN-1993.

XX 30-DEC-1991; 91DE-4143214.

XX 25-JUL-1991; 91DE-4124759.

PR 30-DEC-1991; 91DE-4143214.

XX (BOEF) BOEHRINGER MANNHEIM GMBH.

XX Kaluza B, Riethmuller G, Scheuer W, Weidle U;

XX WPI; 1993-037582/05.

DR N-PSDB; AAQ36614.

XX Synergistic antibody compsn. for use as immunosuppressant -
PT comprises monoclonal anti-CD4 antibodies and monoclonal anti-IL2R
PT alpha- or anti-IL2R beta antibodies
XX Example 1; Page 16; 18pp; German.

XX This sequence is from the heavy chain variable region of an
CC anti-IL2R alpha monoclonal antibody. The antibody is only weakly
CC inhibitory on its own. When anti-IL2R alpha antibodies which are
CC strongly inhibitory on their own are used with anti-CD4 antibodies
CC which are also strongly inhibitory, their immunosuppressive
CC properties are improved; they synergistically inhibit T-helper cell
CC proliferation to effectively inhibit transplant rejection at low
CC doses without significantly reducing the general immune response.
CC No such synergistic effect is observed with the weakly inhibitory
CC MAB M-215.

PS Disclosure; Fig 3C; 74pp; English.

XX The invention provides a monoclonal antibody (Mab) designated 1A7, which
 CC elicits an anti-GD2 (tumor-associated antigen) immunological response in
 CC humans. Mab 1A7 has defined light and heavy chain variable region
 CC sequences. The Mab 1A7 and polypeptides can be used for eliciting an
 CC anti-GD2 immune response. The polypeptides can also be used for detecting
 CC or purifying anti-GD2 antibody. The products can be used for treating GD2
 CC -associated diseases, e.g. melanoma, neuroblastoma, glioma, soft tissue
 CC carcinoma, and small cell carcinoma. They can be used for palliating the
 CC disease or for reducing the risk of recurrence. The present sequence
 CC represents the heavy chain variable region consensus sequence.

XX Sequence 119 AA;

Query Match 76.0%; Score 492; DB 20; Length 119;

Best Local Similarity 78.7%; Pred. No. 1.3e-38;

Matches 96; Conservative 8; Mismatches 14; Indels 4; Gaps 2;

1 QVQLKESGPGLVKPSQSLTICTISGFSLSRYSVHWVRQPPGKGLWLGIMGGNTDYN 60

1 QVQLKESGPGLVKPSQSLTICTISGFSLSRYSVHWVRQPPGKGLWLGIMGGNTDYN 60

61 SALKSRSLTSKDNKSNQVFLKMNLSLTADTAVYYCARK-GEFYYGYDGVFVWGQGLTVTV 119

61 SALKSRSLTSKDNKSNQVFLKMNLSLTADTAVYYCAREXXXYYAMD---YWGQGTSTVTV 117

QY 120 SS 121

DB 118 SS 119

RESULT 13

AAU72838

ID AAU72838 standard; Peptide; 119 AA.

XX AC AAU72838;

XX 26-FEB-2002 (first entry)

XX Anti-NGK2D hybridoma 6H7E7 variable heavy chain CDR1.

XX Human; NGK2D; NGK2D receptor complex; cancer; infectious disease; tumour;
 KW autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;
 KW intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;
 KW prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;
 KW sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAPI0;
 KW helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;
 KW 8G7C10; 6E5A7; 11B2D10x4-7; 8G7C10x4-7; 6E5A7x4-7; P4-2; P4-3; P4-14;
 KW P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;
 KW P53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.

XX Homo sapiens.

XX WO200171005-A2.

XX 27-SEP-2001.

XX 26-MAR-2001; 2001WO-EP03414.

XX 24-MAR-2000; 2000EP-0106467.

XX (KUFE/) KUFER P.

XX Kufer P, Riethmuller G, Lutterbues R, Borschert K, Kischel R;

XX Mayer M, Hofmeister R;

XX WPI; 2002-055119/07.

XX Multifunctional polypeptides comprising binding sites that specifically
 PT recognise extracellular groups of the NGK2D receptor complex and
 PT domains which function as receptors or ligands, useful for treating
 PT cancers and infectious diseases -

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

PI Abrahamson JA, Holmes SD, Jackson JR;
 XX WPI: 2000-387752/33.
 DR N-PSDB; AAA27322.
 XX
 PT Antibodies against receptor for hyaluronic acid mediated motility,
 PT useful for treating or preventing proliferative diseases, e.g. cancer
 PT or cardiovascular disease
 XX
 PS Claim 19; Page 31-32; 39pp; English.
 XX
 CC The present sequence is the protein sequence for the murine 16E10
 CC heavy chain variable region. This forms part of the monoclonal antibody
 CC 16E10, and can be used against the receptor for hyaluronic acid mediated
 CC motility (RHAMM). RHAMM is required for ras-transformation of cells,
 CC which leads to tumour formation, and so the antibody can be used to
 CC treat proliferative disorders such as leukaemias, solid tumour cancers
 CC and metastases including lymphomas, soft tissue, brain, oesophageal,
 CC stomach, pancreatic, liver, lung, bladder, bone, prostate, ovarian,
 CC cervical, uterine, skin, breast, testicular, kidney, head and neck and
 CC colon cancers, chronic inflammatory diseases such as psoriasis,
 CC inflammatory bowel disease and rheumatoid arthritis, proliferative
 CC cardiovascular diseases such as restenosis, proliferative ocular
 CC disorders such as diabetic retinopathy and benign hyperproliferative
 CC diseases such as haemangiomas.
 XX
 SQ Sequence 119 AA;
 Query Match 76.0%; Score 491.5; DB 21; Length 119;
 Best Local Similarity 79.3%; Pred. No. 1.5e-38;
 Matches 96; Conservative 6; Mismatches 16; Indels 3; Gaps 1;
 QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRPPGKGLWGLMGNGGNTDYN 60
 DB 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRPPGKGLWGLMGNGGNTDYN 60
 QY 61 SALKSRSLTSKDKNSKQVFLKMNLSLTADTAVYYCARKGEFFYYGDFVWGQGLTVTVS 120
 DB 61 SALKSRSLTSKDKNSKQVFLKMNLSLTADTAVYYCARKGEFFYYGDFVWGQGLTVTVS 117
 QY 121 S 121
 DB 118 A 118
 RESULT 15
 AAB73623
 AAB73623 standard; Protein; 241 AA.
 AAB73623;
 10-AUG-2001 (first entry)
 Humanised anti-CD40 single chain antibody.
 KW Anti-CD40 antibody; single chain antibody; scFv; humanised;
 KW monoclonal antibody 5H7-derived; autoimmune disease; pernicious anaemia;
 KW Hashimoto's thyroiditis; primary myxoedema thyrotoxicosis;
 KW Addison's disease; insulin-dependent diabetes mellitus; I diabetes;
 KW systemic lupus erythematosus; rheumatoid arthritis; multiple sclerosis;
 KW dermatomyositis; scleroderma; psoriasis; transplant rejection; heart;
 KW liver; pancreatic islet; kidney; lung; bone marrow; drug design;
 KW peripheral blood mononuclear cell activation inhibition; drug screening;
 KW gene therapy.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO200134649-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 08-NOV-2000; 2000WO-US30739.

XX 09-NOV-1999; 99US-0164503.
 PR (CHIR) CHIRON CORP.
 PA
 XX Chu K, Wang C;
 PI
 XX WPI: 2001-329073/34.
 DR N-PSDB; AAH24008.
 XX
 PT New nucleic acid encoding CD40 antagonist, useful for treating
 PT autoimmune disease and transplant rejection, particularly a
 PT single-chain Fv antibody -
 PS Disclosure; Page 47-48; 49pp; English.
 XX
 CC This sequence represents a humanised anti-CD40 single chain antibody
 CC (scFv) derived from the anti-CD40 monoclonal antibody 5H7. The cDNA
 CC encoding it can be incorporated into a gene delivery vector for gene
 CC therapy of autoimmune diseases or transplant rejection. Such gene therapy
 CC vectors containing the polynucleotide of the invention can provide long-
 CC term expression of the anti-CD40 scFv, e.g., for at least 5 years. The
 CC scFv antagonises the action of CD40, signalling through which is
 CC essential for T cell-dependent activation, proliferation and
 CC differentiation of B cells, and which is implicated in B cell survival
 CC and apoptosis. The polynucleotide encoding the scFv is used to prepare
 CC gene therapy vectors for the treatment (or prevention) of autoimmune
 CC diseases, specifically Hashimoto's thyroiditis, primary myxoedema
 CC (type I) diabetes mellitus, systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, dermatomyositis, scleroderma and
 CC especially psoriasis. The scFv polynucleotides can also be used to treat
 CC transplant rejection, especially of heart, liver, pancreatic islet,
 CC kidney, lung or bone marrow tissue. Vectors containing the scFv
 CC polynucleotide may also be used in vitro for inhibiting superantigen
 CC activation of peripheral blood mononuclear cells and to screen for, or
 CC design, potential therapeutic agents.
 XX
 SQ Sequence 241 AA;
 Query Match 75.9%; Score 491; DB 22; Length 241;
 Best Local Similarity 79.3%; Pred. No. 3.6e-38;
 Matches 96; Conservative 8; Mismatches 9; Indels 8; Gaps 2;
 QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRPPGKGLWGLMGNGGNTDYN 60
 DB 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVRPPGKGLWGLMGNGGNTDYN 60
 QY 61 SALKSRSLTSKDKNSKQVFLKMNLSLTADTAVYYCARKGEFFYYGDFVWGQGLTVTVS 120
 DB 61 SALKSRSLTSKDKNSKQVFLKMNLSLTADTAVYYCARKGEFFYYGDFVWGQGLTVTVS 112
 QY 121 S 121
 DB 113 S 113
 RESULT 16
 AAB81973
 ID AAB81973 standard; Protein; 118 AA.
 XX
 AC AAB81973;
 XX
 XX 03-JUL-2001 (first entry)
 XX
 DE Ganglioside GD2 specific antibody related protein SEQ ID NO: 32.
 XX
 KW Ganglioside; GD2; complementation determining region; CDR; antibody;
 KW mouse; cancer.
 XX
 OS Mus musculus.
 XX
 PN WO200123573-A1.

XX PD 05-APR-2001.
 XX PF 29-SEP-2000; 2000WO-JP06773.
 XX PR 30-SEP-1999; 99JP-0278290.
 XX PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX PI Hanai N, Shitara K, Nakamura K, Niwa R;
 XX DR WPI; 2001-266163/27.
 XX CC Human type complementation-determining domain transplanted antibody and
 PT derivatives against ganglioside GD2, useful in diagnosis and therapy of
 PT e.g. tumours, has low antigenicity, little side effects but potent
 PT activity in cancer.
 XX PS Claim 26; Page 115-116; 123pp; Japanese.
 XX CC The present invention describes an antibody, which can react specifically
 CC with ganglioside GD2, and is transplanted with a human type
 CC complementation-determining domain (CDR), or its fragments. The antibody
 CC and its derivatives are useful in diagnosis and therapy of tumours,
 CC particularly cancer diagnosis. The present sequence is a protein
 CC used in the exemplification of the invention.
 XX SQ Sequence 118 AA;
 Query Match 75.7%; Score 489.5; DB 22; Length 118;
 Best Local Similarity 75.2%; Pred. No. 2.2e-38;
 Matches 91; Conservative 14; Mismatches 13; Indels 3; Gaps 1;
 Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWGLMIWGGNTDYN 60
 Db 1 QVQLKESGPGVLVAPSQTLSTICTVSGFSLASYNHWVRQPPGKGLWGLVIWAGGSTNYN 60
 Qy 61 SALKSRLSISKDNSKNOVFLKMNLSLTAAATAVYYCARKGEFYGYDGFVYWGQGLTVTS 120
 Db 61 SALKSRLSISKDNSKNOVFLQMNLSLOTDDTAMYYCAKRSDD--DYSWFAWYWGQGLTVTS 117
 Qy 121 S 121
 Db 118 A 118
 RESULT 17
 AAB81975
 AAB81975 standard; Protein; 137 AA.
 AAB81975;
 XX DT 03-JUL-2001 (first entry)
 XX DE Ganglioside GD2 specific antibody related protein #1.
 XX KW Ganglioside; GD2; complementation determining region; CDR; antibody;
 XX KW mouse; cancer.
 XX OS Mus musculus.
 XX PN WQ200123573-A1.
 XX PD 05-APR-2001.
 XX PF 29-SEP-2000; 2000WO-JP06773.
 XX PR 30-SEP-1999; 99JP-0278290.
 XX PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX PI Hanai N, Shitara K, Nakamura K, Niwa R;
 XX DR WPI; 2001-266163/27.
 XX CC Human type complementation-determining domain transplanted antibody and
 PT derivatives against ganglioside GD2, useful in diagnosis and therapy of
 PT e.g. tumours, has low antigenicity, little side effects but potent
 PT activity in cancer.
 XX PS Claim 26; Page 115-116; 123pp; Japanese.
 XX CC The present invention describes an antibody, which can react specifically
 CC with ganglioside GD2, and is transplanted with a human type
 CC complementation-determining domain (CDR), or its fragments. The antibody
 CC and its derivatives are useful in diagnosis and therapy of tumours,
 CC particularly cancer diagnosis. The present sequence is a protein
 CC used in the exemplification of the invention.
 XX SQ Sequence 118 AA;
 Query Match 75.7%; Score 489.5; DB 22; Length 118;
 Best Local Similarity 75.2%; Pred. No. 2.2e-38;
 Matches 91; Conservative 14; Mismatches 13; Indels 3; Gaps 1;
 Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWGLMIWGGNTDYN 60
 Db 1 QVQLKESGPGVLVAPSQTLSTICTVSGFSLASYNHWVRQPPGKGLWGLVIWAGGSTNYN 60
 Qy 61 SALKSRLSISKDNSKNOVFLKMNLSLTAAATAVYYCARKGEFYGYDGFVYWGQGLTVTS 120
 Db 61 SALKSRLSISKDNSKNOVFLQMNLSLOTDDTAMYYCAKRSDD--DYSWFAWYWGQGLTVTS 117
 Qy 121 S 121
 Db 118 A 118

DR WPI; 2001-266163/27.
 DR N-PSDB; AAF86854.
 XX Human type complementation-determining domain transplanted antibody and
 PT derivatives against ganglioside GD2, useful in diagnosis and therapy of
 PT e.g. tumours, has low antigenicity, little side effects but potent
 PT activity in cancer.
 XX PS Example 2; Page 96-97; 123pp; Japanese.
 XX CC The present invention describes an antibody, which can react specifically
 CC with ganglioside GD2, and is transplanted with a human type
 CC complementation-determining domain (CDR), or its fragments. The antibody
 CC and its derivatives are useful in diagnosis and therapy of tumours,
 CC particularly cancer diagnosis. The present sequence is a protein
 CC used in the exemplification of the invention.
 XX SQ Sequence 137 AA;
 Query Match 75.7%; Score 489.5; DB 22; Length 137;
 Best Local Similarity 75.2%; Pred. No. 2.6e-38;
 Matches 91; Conservative 14; Mismatches 13; Indels 3; Gaps 1;
 Qy 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWGLMIWGGNTDYN 60
 Db 20 QVQLKESGPGVLVAPSQTLSTICTVSGFSLASYNHWVRQPPGKGLWGLVIWAGGSTNYN 79
 Qy 61 SALKSRLSISKDNSKNOVFLKMNLSLTAAATAVYYCARKGEFYGYDGFVYWGQGLTVTS 120
 Db 80 SALKSRLSISKDNSKNOVFLQMNLSLOTDDTAMYYCAKRSDD--DYSWFAWYWGQGLTVTS 136
 Qy 121 S 121
 Db 137 A 137
 RESULT 18
 AAP70991
 AAP70991 standard; Protein; 135 AA.
 XX AC AAP70991;
 XX DT 20-JUN-1991 (first entry)
 XX DE Sequence of the heavy chain variable region of the D1.3
 DE antibody.
 XX KW Recombinant altered antibody; humanised monoclonal antibody.
 XX FH Key Location/Qualifiers
 FT Peptide 1..19
 FT Region /label= Signal
 FT 49..51
 FT Region /note= "residues which contact lysozyme"
 FT 71..73
 FT Region /note= "residues which contact lysozyme"
 FT 118..121
 FT Region /note= "residues which contact lysozyme"
 FT 50..54
 FT Region /label= CDR 1
 FT 69..84
 FT Region /label= CDR 2
 FT 117..124
 FT Region /label= CDR 3
 XX EP239400-A.
 XX PD 30-SEP-1987.
 XX PF 26-MAR-1987;
 XX PR 27-MAR-1986;
 XX PR 26-MAR-1987;
 XX PR 87EP-0302620.
 XX PR 86GB-0007679.
 XX PR 87GB-0007252.


```
XX (WINT/) WINTER G P.
XX Winter GP;
XX WPI: 1987-272811/39.
XX N-PSDB; AAN72111.
XX
XX Recombinant altered antibodies - having complementarity
XX PT determining regions replaced with those from antibody of
XX PT different specificity
XX PS Example: Fig 7; 4lpp; English.
XX
XX The method of the invention is used for "humanising" non-human
XX CC monoclonal antibodies (MABs) eg CDRs from mouse MAB can be
XX CC partially or totally grafted into the framework regions of a human
XX CC MAB, which is then produced in quantity by a suitable cell line.
XX CC CDS- complementarity determining regions. The antibody has
XX CC DI.3 (H2K2) has two heavy chains of the mouse IgG1 class (H) and
XX CC two kappa light chains (K).
XX
XX Sequence 135 AA:
XX Query Match 75.5%; Score 488.5; DB 8; Length 135;
XX Best Local Similarity 79.3%; Pred. No. 3.2e-38;
XX Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;
XX
XX 1 QVQLKESGPGLVKPSQTLITCTISGFSLSRYSVHVRQPPGKGLWLGMIWGNGNTDYN 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 20 QVQLKESGPGLVAPSQSLITCTVSGFSLTGYVNVVRQPPGKGLWLGMIWGNGNTDYN 79
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 61 SALKSRLSISKDKNQVFLKMSLTAADTAVYVCARKGEFYGYDGFVYWGQGLTIVTS 120
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 80 SALKSRLSISKDKNQVFLKMSLTDITDITARYCARERD--YRLD---YWGQGLTIVTS 134
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 121 S 121
XX 135 S 135
XX
XX RESULT 19
XX AAR56231
XX ID AAR56231 standard; Protein; 211 AA.
XX AC AAR56231;
XX
XX 04-JAN-1995 (first entry)
XX 161-107/113-60/h1-116 DI.3 V-min intact heavy chain.
XX
XX MCP603; V-min; synthetic gene; DI.3; antibody; immunotoxins; tumour;
XX KW intracellular expression; E. coli; protease deficient; treatment;
XX KW ligand-binding variable domain; framework region; beta-barrel;
XX KW beta-sheet; linker; complementarity determining regions; radioimaging;
XX KW produg therapy; diagnosis.
XX
XX Synthetic.
XX OS
XX Key Location/Qualifiers
XX FH Region 1..47
XX FT /note= "161-107"
XX FT 48..95
XX FT /note= "113-160"
XX FT 96..211
XX FT /note= "h1-h116"
XX
XX WO9412625-A.
XX PN
XX XX
XX PD 09-JUN-1994.
XX XX
XX PF 19-NOV-1993; 93WO-CB02375.
XX XX
```

```
PR 23-NOV-1992; 92GB-0024588.
XX 31-DEC-1992; 92GB-0027189.
XX (ZENE ) ZENECA LTD.
XX
XX Slater AM, Timms D;
XX
XX WPI: 1994-200254/24.
XX
XX New ligand variable binding domains of reduced size - contain
XX PT complementarity determining regions and framework region to
XX PT provide correct orientation, also related DNA, expression vectors
XX PT etc., useful for radio-imaging and tumour treatment
XX
XX Disclosure: Fig 42; 187pp; English.
XX
XX The sequences given in AAR56229-36 show various V-min constructs. V-min
XX is a ligand-binding variable domain which comprises a framework region
XX of a cyclically permuted central beta-barrel, outer beta-sheet
XX segments, and linker segments, and complementarity determining regions.
XX V-min, when conjugated, are useful in radioimaging, as immunotoxins
XX and in antibody directed enzyme produg therapy. ie. for diagnosis or
XX treatment of tumours. Compared to ordinary antibodies, V-min are
XX significantly smaller, as few as 160 amino acids, and so may show
XX improved tissue penetration and faster clearance from the body. Small
XX size should facilitate conjugation and oligomerisation. V-min can be
XX expressed as a continuous polypeptide from a single gene.
XX
XX Sequence 211 AA:
XX Query Match 75.5%; Score 488.5; DB 15; Length 211;
XX Best Local Similarity 79.3%; Pred. No. 5.3e-38;
XX Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;
XX
XX 1 QVQLKESGPGLVKPSQTLITCTISGFSLSRYSVHVRQPPGKGLWLGMIWGNGNTDYN 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 96 QVQLKESGPGLVAPSQSLITCTVSGFSLTGYVNVVRQPPGKGLWLGMIWGNGNTDYN 155
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 61 SALKSRLSISKDKNQVFLKMSLTAADTAVYVCARKGEFYGYDGFVYWGQGLTIVTS 120
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 156 SALKSRLSISKDKNQVFLKMSLTDITDITARYCARERD--YRLD---YWGQGLTIVTS 210
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 121 S 121
XX 211 S 211
XX
XX RESULT 20
XX AAR06355
XX ID AAR06355 standard; peptide; 127 AA.
XX
XX AAR06355;
XX
XX 25-SEP-1990 (first entry)
XX
XX Peptide corresponding to PAC1 murine monoclonal antibody.
XX
XX Fibrinogen binding; PAC1 murine; platelet aggregation;
XX KW thrombosis; cancer metastasis; glycoprotein IIB-IIIA complex;
XX KW heavy chain; hypervariable region.
XX
XX Synthetic.
XX OS
XX Key Location/Qualifiers
XX FH Peptide 1..103
XX FT /label=(X1)n
XX FT /note="n=0 or 1"
XX FT 96..103
XX FT /label=(X1)n
XX FT /note="n=0 or 1"
XX FT 107..127
XX FT /label=(X2)m
XX FT /note="m=0 or 1"
```

FT Peptide 107..116
FT /label=(X2)m
FT /note="N=0 Or 1"

PN EP368486-A.

PD 16-MAY-1990.

PF 17-OCT-1989; 89EP-0310620.

PR 10-NOV-1989; 88US-0269583.

XX (MERI) MERCK AND CO INC.

PA (UYPE-) UNIVERSITY OF PENNSYLVANIA.

XX Shattil SJ, Taub RA, Friedman PA;

PI WPI; 1990-149427/20.

DR Peptide(s) corresponding to PAC1 murine monoclonal antibody -

XX used for inhibiting fibrinogen binding, platelet aggregation,

XX thrombosis and/or cancer metastasis.

PS Claim 8; Page 11; 17pp; English.

XX The peptide has the general formula: R1-(X1)n-R2 where R1 is a

CC (un)protected N-terminal gp and R2 is a (un)protected C-terminal gp.

CC The peptide shown is the longest given but truncated versions esp. as

CC indicated in the feature table are also used. The peptides inhibit the

CC binding of fibrinogen to blood platelets, platelet aggregation and/or

CC binding of cpds. or substances, partic. proteins, to the glycoprotein

CC IIB-IIIa complex found in blood platelet membranes. They are therefore

CC useful for preventing thrombosis and cancer metastasis. The -R2D-

CC sequence is present in one of the heavy chain hypervariable regions of

CC PAC1 murine monoclonal antibody.

XX Sequence 127 AA;

Query Match 75.4%; Score 488; DB 11; Length 127;

Best Local Similarity 72.4%; Pred. No. 3 3e-38;

Matches 92; Conservative 11; Mismatches 18; Indels 6; Gaps 1;

OY 1 QVQLKESGFLVKKPQSTISITCTISGFSLSRYSVHVRPPGKLEWLGIMGGNTDYN 60

Db 1 QVQLKESGFLVKKPQSTISITCTISGFSLSRYSVHVRPPGKLEWLGIMGGNTDYN 60

OY 61 SALKSRSLTSKDNKSNQVFLKNSLTAAATVAVYCCARKGEFYVGGQ 114

61 AAFISRLTSKDNKSNQVFLKNSLTAAATVAVYCCARKGEFYVGGQ 120

OY 115 TLTVTSS 121

Db 121 TSVTVSS 127

RESULT 21

AAU07513

XX AAU07513 standard; Peptide; 116 AA.

AC AAU07513;

XX 24-OCT-2001 (first entry)

XX Antibody scFv8 heavy chain variable region, mutant D1.3.

XX Antimicrobial; antiviral; cytostatic; immunomodulatory;

KW antibody; gene therapy; HIV; human immunodeficiency virus; tumour;

KW metabolic disorder; immune disorder; auto-immune disorder; lysozyme;

KW antibody heavy chain variable region; scFv8; mutant D1.3; mutein.

XX Synthetic.

XX Key Location/Qualifiers

XX Key

FT Misc-difference 9..10 /note="Wild-type Gly-Asp substituted by Pro-Gly"

FT Misc-difference 13 /note="Wild-type Gln substituted by Ala"

FT Misc-difference 15..16 /note="Wild-type Gly-Gly substituted by Ser-Gln"

FT Misc-difference 19..21 /note="Wild-type Lys-Leu-Ser substituted by Ser-Ile-Thr"

FT Misc-difference 23..24 /note="Wild-type Ala-Ala substituted by Thr-Val"

FT Misc-difference 28..31 /note="Wild-type Thr-Phe-Ser-Ser substituted by Ser-Leu-Thr-Gly"

FT Misc-difference 34..35 /note="Wild-type Met-Ser substituted by Val-Asn"

FT Misc-difference 40 /note="Wild-type Thr substituted by Pro"

FT Misc-difference 42 /note="Wild-type Asp substituted by Gly"

FT Misc-difference 44 /note="Wild-type Arg substituted by Gly"

FT Misc-difference 47..50 /note="Wild-type Leu-Val-Ala-Thr substituted by Trp-Leu-Gly-Met"

FT Misc-difference 52..54 /note="Wild-type Asn-Ser-Asn substituted by Trp-Gly-Asp"

FT Misc-difference 56 /note="Wild-type Ser substituted by Asn"

FT Misc-difference 58 /note="Wild-type Phe substituted by Asp"

FT Misc-difference 60..63 /note="Wild-type Pro-Asp-Ser-Val substituted by Asn-Ser-Ala-Leu"

FT Misc-difference 65 /note="Wild-type Gly substituted by Ser"

FT Misc-difference 67..68 /note="Wild-type Phe-Thr substituted by Leu-Ser"

FT Misc-difference 71 /note="Wild-type Arg substituted by Lys"

FT Misc-difference 74 /note="Wild-type Ala substituted by Ser"

FT Misc-difference 76..79 /note="Wild-type Asn-Thr-Leu-Tyr substituted by Ser-Gln-Val-Phe"

FT Misc-difference 81 /note="Wild-type Gln substituted by Lys"

FT Misc-difference 83 /note="Wild-type Ser substituted by Asn"

FT Misc-difference 86..88 /note="Wild-type Lys-Ser-Glu substituted by His-Thr-Asp"

FT Misc-difference 92 /note="Wild-type Met substituted by Arg"

FT Misc-difference 98 /note="Wild-type Arg substituted by Glu"

FT Misc-difference 100 /note="Wild-type Asn substituted by Asp"

FT Misc-difference 102..103 /note="Wild-type Pro-Tyr-Tyr-Tyr-Gly-Ser-Arg-Tyr-Phe substituted by Arg-Leu"

XX WO200149713-A2.

XX 12-JUL-2001.

XX 29-DEC-2000; 2000WO-IT00554.

XX 30-DEC-1999; 99IT-RM00803.

XX (CNEN) ENEA ENTE NUOVE TECNOLOGIE ENERGIA.

PA (CONS-) SOC CONSORTILE METAPONTUM AGROBIOS SRL.

XX Benvenuto E, Franconi R, Desiderio A, Tavladoraki P;
 PI WPI; 2001-502555/55.
 XX Peptides which are able to confer stability and solubility to an
 PT antibody comprising these peptides, useful for treating pathologies
 PT (e.g. tumour) associated with accumulation of a molecule inside or
 PT outside a human, or animal cell -
 XX Disclosure; Fig 4; 109pp; English.
 PS
 XX The invention relates to peptides which are able to confer stability and
 CC solubility to an antibody comprising these peptides. The peptides
 CC are especially H-FR1, H-FR2, H-FR3, H-FR4, L-FR1, L-FR2, L-FR3 or L-FR4
 CC present within a variable region of an antibody which makes the antibody
 CC soluble and stable in cytoplasm. Peptides having the sequences of HFR1 to
 CC H-FR4 are present within the variable region of the heavy chain of an
 CC antibody, covalently linked to the H-CDR1, H-CDR2, H-CDR3 in the order
 CC (H-FR1)-(H-CDR1)-(H-FR2)-(H-CDR2)-(H-FR3)-(H-CDR3)-(H-CDR4) and
 CC peptides having the sequences of L-FR1 to L-FR4 are present within
 CC the variable region of the light chain of an antibody, covalently
 CC linked to the L-CDR1, L-CDR2, L-CDR3 in the order (L-FR1)-(L-CDR1)-
 CC (L-FR2)-(L-CDR2)-(L-FR3)-(L-CDR3)-(L-CDR4). The antibodies and
 CC polynucleotides (e.g. by gene therapy) are useful for the manufacture
 CC of a medicament for the treatment of pathologies associated with
 CC accumulation of a molecule inside or outside a human, animal cell
 CC or plant cell. The pathologies are infectious (e.g. viral infections such
 CC as HIV, human immunodeficiency virus, infections), tumour, metabolic and
 CC immune (especially auto-immune) pathologies. The present sequence
 CC represents the antibody scfv8 heavy chain variable region, mutant D1.3.
 XX
 SQ Sequence 116 AA;
 Query Match 75.3%; Score 487.5; DB 22; Length 116;
 Best Local Similarity 79.3%; Pred. No. 3.3e-38;
 Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;
 QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVPPKGLGLEWLGMIWGNTDYN 60
 DB 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVPPKGLGLEWLGMIWGNTDYN 60
 QY 61 SALKSRLSISKDNKQVFLKMSLTAAADTAVYVCARKGEFYGYDGFFVWGQGLTVTS 120
 DB 61 SALKSRLSISKDNKQVFLKMSLTAAADTAVYVCARKGEFYGYDGFFVWGQGLTVTS 120
 QY 121 S 121
 116 S 116
 RESULT 22
 AAY22432
 ID AAY22432 standard; protein; 120 AA.
 AC AAY22432;
 XX 28-SEP-1999 (first entry)
 XX TM27 antibody VH chain mutant V67L, T68S, M69I, L70S, T73N.
 XX IgG antibody; TM27; variable chain; immunoglobulin G; medical therapy;
 KW autoimmune disease; multiple sclerosis; antibody storage; VH; mutein;
 XX heavy chain.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 67
 FT /label= V67L
 FT Misc-difference 68
 FT /label= T68S
 FT Misc-difference 69

FT Misc-difference 70
 FT /label= L70S
 FT Misc-difference 73
 FT /label= T73N
 XX WO9937329-A1.
 XX 29-JUL-1999.
 XX 15-JAN-1999; 99WO-SE00049.
 XX 09-MAR-1998; 98SE-0000766.
 XX 22-JAN-1998; 98SE-0000170.
 XX (ASTR) ASTRA AB.
 XX Flink O, Petren S;
 XX WPI; 1999-458611/38.
 PT Isotonic pharmaceutical antibody formulations comprising a citrate
 PT buffer, have improved storage
 PS Claim 12; Page 23-24; 30pp; English.
 XX This sequence is a mutant of the variable heavy (VH) chain of the
 CC antibody TM27. TM27 is an immunoglobulin G (IgG) antibody, contained in
 CC the isotonic pharmaceutical formulation of the invention, along with a
 CC citrate buffer at a physiologically acceptable pH. The formulations are
 CC useful in medical therapy, especially for treatment of autoimmune
 CC disease, and particularly, therapy of multiple sclerosis (using the TM27
 CC antibody heavy or light chains). The formulation is useful for improving
 CC the storage of an antibody. The antibody formulations are simple and
 CC effective. The formulations are stable and have improved storage
 CC properties. The formulation is simpler than prior art formulations.
 CC Existing antibody formulations require the use both of a stabiliser and a
 CC buffer. The present invention formulations are stabilized only by citrate
 CC buffer in a saline solution at a physiologically preferable pH. The
 CC avoidance of low pH prevents undesirable reaction as the site of
 CC injection. The formulation does not use ovalbumin for stabilisation,
 CC hence avoiding an allergic response to ovalbumin. Also the formulation
 CC does not require lyophilisation which is an expensive process and also
 CC requires the formulation to be resuspended prior to administration.
 XX
 SQ Sequence 120 AA;
 Query Match 75.3%; Score 487.5; DB 20; Length 120;
 Best Local Similarity 76.4%; Pred. No. 3.5e-38;
 Matches 94; Conservative 12; Mismatches 12; Indels 5; Gaps 2;
 QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVPPKGLGLEWLGMIWGNTDYN 60
 DB 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVPPKGLGLEWLGMIWGNTDYN 60
 QY 61 SALKSRLSISKDNKQVFLKMSLTAAADTAVYVCARK--GEFYGYDGFFVWGQGLTVT 118
 DB 61 SALKSRLSISKDNKQVFLKMSLTAAADTAVYVCARKDRVTATLYAMD---YWGQGLTVT 117
 QY 119 VSS 121
 118 VSS 120
 RESULT 23
 AAW35561
 ID AAW35561 standard; Protein: 239 AA.
 XX AAW35561;
 XX 24-APR-1998 (first entry)
 XX EcoRI-HindIII insert of pUR.4124 protein sequence.

```

XX Multivalent antigen binding protein; variable domain; binding site;
KW diagnosis; therapy; medicine; immunoassay; target; PUR4124.
XX Synthetic.
OS
XX Key Location/Qualifiers
FH 1..108
FT Protein /note= "Mature protein VLlys"
FT Peptide /note= "109..123"
FT Protein /note= "linker peptide"
FT 124..239
FT /note= "Mature protein VHlys"
XX WO9738102-A1.
XX
XX 16-OCT-1997.
XX
XX 26-MAR-1997; 97WO-EP01609.
XX
XX 04-APR-1996; 96EP-0302412.
XX (UNIL ) UNILEVER NV.
XX (UNIL ) UNILEVER PLC.
XX Davis PJ, Van DER LOGT CPE, Verhoeven ME;
XX WPI: 1997-512719/47.
XX N-PSDB; AAT94606.
XX
XX Multivalent antigen binding protein - comprises antibody variable
XX domains, linked by their association, where each associated variable
XX domain pair forms an antigen binding site
XX
XX Example 3; Pages 42-43; 100pp; English.
XX
XX This is the protein sequence of EcoRI-HindIII insert of PUR4124. The
XX corresponding DNA encodes VHlys-linker-VLlys sequences. This is used in
XX the construction of a novel multivalent antigen binding protein. This
XX novel protein comprises 2 polypeptides, each comprising, in series, 3 or
XX more variable domains of an antibody heavy or light chain respectively,
XX which are linked by association of the respective heavy and light chain
XX variable domains, where each associated variable domain pair forms an
XX antigen binding site. The multivalent antigen binding protein can be used
XX in medicine, diagnosis, therapy, an immunoassay method or for
XX purification. It may be used in targeting a tumour cell with natural
XX killer cells and cytotoxic agents. It may be used to target cell killing
XX enzymes, e.g. oxidases and peroxidases, to a species with which is an
XX antigenic component of a dental plaque, such as S. sanguis or S. mutans.
XX The antigen binding protein may conveniently be purified straight from
XX the supernatant using conventional purification techniques. As the
XX protein is self-assembling, there is no need to purify individual
XX subunits prior to coupling as in existing techniques.
XX
XX Sequence 239 AA;
XX Query Match 75.3%; Score 487.5; DB 18; Length 239;
XX Best Local Similarity 79.3%; Pred. No. 7.5e-38;
XX Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;
XX
XX 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWRQPPGKGLWGLMIGWGNTDYN 60
XX 124 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWRQPPGKGLWGLMIGWGNTDYN 183
XX
XX 61 SALKSRSLISKDNKSNQVFLKMNLSLTADTAVYVCARKGEFFYYGDFGVYWGQGLTVTVS 120
XX 184 SALKSRSLISKDNKSNQVFLKMNLSLTADTAVYVCARKGEFFYYGDFGVYWGQGLTVTVS 238
XX
XX 121 S 121
XX
XX 239 S 239
XX

```

```

RESULT 24
AAW09813
ID AAW09813 standard; Protein; 239 AA.
XX
XX AC AAW09813;
XX
XX 14-JUL-1997 (first entry)
XX
XX VLlys-linker-VHlys polypeptide.
XX
XX Bispecific antibody; bivalent antibody; double head;
XX antibody engineering: GOSA; glucose oxidase; Streptococcus sanguis.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
FH 1..106
FT Protein /label= VLlys
FT Peptide 109..123
FT /label= Linker
FT Protein 124..239
FT /label= VHlys
XX
XX WO9714719-A1.
XX
XX 24-APR-1997.
XX
XX 14-AUG-1996; 96WO-EP03605.
XX
XX 16-OCT-1995; 95EP-0307332.
XX (UNIL ) UNILEVER NV.
XX (UNIL ) UNILEVER PLC.
XX
XX Davis PJ, Van Der Logt CPE, Verhoeijen ME, Wilson S;
XX
XX WPI: 1997-245049/22.
XX N-PSDB; AAT66143.
XX
XX Bispecific or bivalent antibody fragment analogues comprise complex
XX of two polypeptide chains containing respectively two heavy and two
XX light chain variable domains - useful therapeutically, in
XX immunoassays, for purification etc.
XX
XX Example 1; Page 42-43; 121pp; English.
XX
XX A polypeptide (AAW09813) encoded by the EcoRI-HindIII insert (AAT66143)
XX of plasmid PUR.4124 comprises a VL domain linked to a VH domain via
XX a flexible linker. Novel polypeptides (see also AAW09814-20) can be
XX utilised in novel bispecific or bivalent antibodies that consist of
XX a binding complex of two polypeptide chains, one with two VH
XX domains in series and the other with two VL domains in series, the
XX two domains being connected either directly or via a linker peptide,
XX and the complex contg. two pairs of variable domains (VH-A/VL-A and
XX VH-B/VL-B). The two variable domains may be different, producing
XX a bispecific construct or A and B are the same and the construct is
XX bivalent. The novel antibody fragment analogues are useful in
XX immunoassays, purification and therapeutically.
XX
XX Sequence 239 AA;
XX Query Match 75.3%; Score 487.5; DB 18; Length 239;
XX Best Local Similarity 79.3%; Pred. No. 7.5e-38;
XX Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;
XX
XX 1 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWRQPPGKGLWGLMIGWGNTDYN 60
XX 124 QVQLKESGPGLVKPSQTLISITCTISGFSLSRYSVHWRQPPGKGLWGLMIGWGNTDYN 183
XX
XX 61 SALKSRSLISKDNKSNQVFLKMNLSLTADTAVYVCARKGEFFYYGDFGVYWGQGLTVTVS 120
XX 184 SALKSRSLISKDNKSNQVFLKMNLSLTADTAVYVCARKGEFFYYGDFGVYWGQGLTVTVS 238
XX

```

QY 121 S 121
DB 239 S 239

RESULT 25
AAR21261
ID AAR21261 standard; Protein: 241 AA.

AC AAR21261;
XX 21-MAY-1992 (first entry)

DE VHD1.3-HUCH1 region of Fab D1.3 in pUC19.

FW Fd; bacteriophage; gene III; filamentous; phagemid; capsid; coat;
KW pilus; g3p; binding; adsorption; gene VIII; diverse repertoire;
KW specific binding pairs; replicable genetic display package; Ig;
KW immunoglobulin; antibody; heavy chain; light chain; variable.

Synthetic.

Key Location/Qualifiers
FT Peptide 1..22
FT Region /label= PelB_leader
FT 23..138
FT /label= VHD1.3
FT 139..241
FT /label= HUCH1

PN W09201047-A.

XX 23-JAN-1992.

XX 10-JUL-1991; 91WO-GB01134.

XX 15-MAY-1991; 91GB-0010549.

XX 10-JUL-1990; 90GB-0015198.

XX 19-OCT-1990; 90GB-0022845.

XX 12-NOV-1990; 90GB-0024503.

XX 06-MAR-1991; 91GB-0004744.

XX (CAMP-) CAMBRIDGE ANTIBODY.

PA (MEDI-) MED RES COUNCIL.

XX McCafferty J, Pope AR, Johnson KS, Hoogenboom HRJ, Griffiths AD;
PI Jackson RH, Holliger KP, Marks JD, Clackson TP, Chiswell DJ;
Winter GP, Bonfert TP;

DR WPI; 1992-056862/07.

XX N-PSDB: AAQ21097.

XX Producing members of specific binding pairs - by expression in
PT recombinant host cells with a secreting replicable genetic
PT display package.

XX Example 7: Fig 10; 209pp; English.

XX The plasmid was the starting point for the prepn. of a pAb that
CC expresses a Fab-like fragment on its surface. (The antibody D1.3
CC is directed against hen egg lysozyme). The coding sequence
CC are preceded by the pelB signal peptide sequence (Ward ES et al,
CC 1989., Nature 341 544-546 (1989)). The sequence encoding the VH
CC and CH1 regions shown here was amplified and ligated into fdCAT2
CC (see AAQ23463) which contains the gene III sequence to construct
CC fdCAT2VHCH1 D1.3. The heavy chain was deleted from Fab D1.3 in
CC pUC19 and the 2.7 kb fragment contg. the light chain was purified
CC and self-ligated to prepare LCD1.3 DHC. Supernatant contg. phage
CC particles from cells transformed with fd CAT2VHCH1 D1.3 were
CC cultured with LCD1.3 DHC cells and phage from the resulting colonies
CC were assayed for ability to bind to lysozyme. Results showed that
CC a functional Fab fragment was produced by an association of the free
CC light chain (see AAR22565) with the VHCH1 fragment (shown here) fused

CC to gene III and expressed on the surface of the pAb.
CC See also AAR21260-307, 309-311; AAR22450, 565-581.

SQ Sequence 241 AA;

Query Match 75.3%; Score 487.5; DB 13; Length 241;
Best Local Similarity 79.3%; Pred. No. 7.6e-38;
Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGNTDYN 60
DB 23 QVQLQESGPGLVAPSQSLSTCTISGFSLSITCTVSGFSLTGYNWVRQPPGKGLEWLGMIWGNTDYN 82

QY 61 SALKSRLSISKDNKSNQVFLKMSLTAAATVAYYCARKGEFFYYGYDGFVYWGQGTFLVTVS 120
DB 83 SALKSRLSISKDNKSKSQVFLKMSLTHTDDTARYYCAREDD--YRLD---YWGQGTFLVTVS 137

QY 121 S 121
DB 138 S 138

Search completed: June 23, 2003, 14:01:42
Job time : 22.5231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 18.0709 Seconds
(without alignments)
825.861 Million cell updates/sec

Title: US-10-056-052A-18
Perfect score: 583
Sequence: 1 DIVWTQSPDSLAVSLGERVT.....YCHQYLSSYTFGGTKLEIK 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

al number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	551	94.5	112	15	AA1954933
2	540	92.6	263	20	AA190226
3	540	92.6	268	20	AA190228
4	540	92.6	268	20	AA190222
5	540	92.6	273	20	AA190224
6	540	92.6	556	20	AA190218
7	540	92.6	580	20	AA190217
8	527	90.4	113	17	AA192215
9	527	90.4	113	18	AA192215
10	526	90.2	113	22	AA192215

11	525.5	90.1	113	15	AA190313
12	524.5	90.0	120	22	AA190313
13	521	89.4	112	22	AA190313
14	520.5	89.3	113	15	AA190314
15	520.5	89.3	155	18	AA190314
16	520.5	89.3	155	20	AA190314
17	520.5	89.3	342	18	AA190314
18	520.5	89.3	342	18	AA190314
19	520.5	89.3	495	18	AA190314
20	520.5	89.3	495	18	AA190314
21	516.5	88.6	137	21	AA190314
22	515	88.3	241	13	AA190314
23	513	88.0	259	21	AA190314
24	512.5	87.9	133	9	AA190314
25	512.5	87.9	134	20	AA190314
26	512.5	87.9	171	14	AA190314
27	512.5	87.9	171	20	AA190314
28	512.5	87.9	171	20	AA190314
29	512.5	87.9	240	22	AA190314
30	512.5	87.9	251	22	AA190314
31	512.5	87.9	258	22	AA190314
32	512.5	87.9	274	14	AA190314
33	512.5	87.9	274	20	AA190314
34	512.5	87.9	274	20	AA190314
35	512.5	87.9	284	14	AA190314
36	512.5	87.9	284	20	AA190314
37	512.5	87.9	284	20	AA190314
38	511	87.7	114	15	AA190314
39	510.5	87.6	113	21	AA190314
40	510.5	87.6	119	23	AA190314
41	510	87.5	141	20	AA190314
42	508.5	87.2	137	18	AA190314
43	508.5	87.2	137	18	AA190314
44	507.5	87.0	113	21	AA190314
45	507.5	87.0	114	14	AA190314
46	507.5	87.0	114	23	AA190314
47	507.5	87.0	133	15	AA190314
48	507.5	87.0	264	23	AA190314
49	506.5	86.9	260	20	AA190314
50	506.5	86.9	260	20	AA190314
51	506	86.8	113	17	AA190314
52	505.5	86.7	111	15	AA190314
53	505.5	86.7	113	21	AA190314
54	505.5	86.7	113	21	AA190314
55	505.5	86.7	134	22	AA190314
56	505.5	86.7	285	15	AA190314
57	504.5	86.5	113	15	AA190314
58	504.5	86.5	133	15	AA190314
59	503.5	86.4	113	20	AA190314
60	503.5	86.4	115	18	AA190314
61	502.5	86.2	113	21	AA190314
62	502.5	86.2	113	21	AA190314
63	502.5	86.2	113	21	AA190314
64	502.5	86.2	113	21	AA190314
65	502.5	86.2	135	21	AA190314
66	502.5	86.2	139	17	AA190314
67	501.5	86.0	113	21	AA190314
68	501.5	86.0	113	21	AA190314
69	501	85.9	115	18	AA190314
70	501	85.9	134	18	AA190314
71	500.5	85.8	110	13	AA190314
72	499.5	85.7	113	20	AA190314
73	499.5	85.7	113	21	AA190314
74	499.5	85.7	240	20	AA190314
75	499.5	85.7	432	23	AA190314
76	499.5	85.7	480	23	AA190314
77	499.5	85.7	601	23	AA190314
78	499.5	85.7	614	23	AA190314
79	499.5	85.7	658	23	AA190314
80	498	85.4	112	15	AA190314
81	498	85.4	301	18	AA190314
82	498	85.4	301	20	AA190314
83	498	85.4	301	22	AA190314

Humanised light ch
Amino acid sequenc
5A8 VL. Synthetic
Humanised light ch
Kappa light chain
Human variable kap
TLHL amino acid se
Growth factor CATA
CATAB-TEV aminoc
Humanised antibody
Vector pMD1007.
Antiviral scfv-ant
V region of L chai
Human Hum4 VL Clai
Sequence of signal
Amino acid sequenc
Plasmid pATDFLAG F
Human immunoglobul
Single chain antib
Single chain antib
Sequence of Hum4 V
Amino acid sequenc
Human Hum4L-CC49 V
Sequence of pSC49F
Amino acid sequenc
Plasmid pSC49FLAG
Monoclonal antibody
260F9 hybridoma VL
ebvHlgM MSt19D10 1
Human monoclonal a
Mouse MAb 15 light
Chimeric MAb 15 PC
Amino acid sequenc
MAb GAH variable r
Human Len kappa 11
Sequence of Hum4VL
Human ovarian anti
Amino acid sequenc
Human SCFV1 protei
Humanised LL2 MAB
Light chain variab
Amino acid sequenc
Amino acid sequenc
Anti-CA125 bifunct
Sequence of plasm
MAb NFS2 light cha
Sequence of the 11
Antibody F19 human
Human Ab light cha
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc
Variable kappa cha
Xenograft antibody
Light chain variab
Antibody F19 human
Amino acid sequenc
Human reshaped F19
Antibody-cytokine
TNF-selectokine pr
Antibody-cytokine
TNF-selectokine pr
Fc receptor huma
Single chain, huma
Multispecific sing
Single chain human

Single chain human
Anti-Fc gamma rece
Single chain anti-
H22-anti-CEA antib
Bispecific single
Murine kappa chain
Mouse HMFGL light
Human trkC antibod
Consensus contrace
Sequence encoded b
Light chain variab
Amino acid sequenc
Human IGFAM-17 imm
Sequence of single
Anti-HGF receptor
Amino acid sequenc

84 498 85.4 301 22 AAB61959
85 498 85.4 352 20 AAY06272
86 498 85.4 553 18 AAW11508
87 498 85.4 553 20 AAW73223
88 498 85.4 553 22 AAB54455
89 498 85.4 553 22 AAB61960
90 496.5 85.2 114 13 AAR22419
91 496.5 85.2 114 15 AAR3803
92 496.5 85.2 116 23 AAU81281
93 496.5 85.2 133 21 AAY83647
94 496.5 85.2 134 21 AAR38317
95 495.5 85.0 113 21 AAR25407
96 494.5 84.8 113 21 AAB18867
97 494.5 84.8 244 21 AAY96305
98 494.5 84.8 260 14 AAR38318
99 490.5 84.1 220 18 AAW07528
100 490 84.0 132 21 AAB07964

ALIGNMENTS

RESULT 1
AAR54933 standard; peptide: 112 AA.

XX AAR54933;
XX 19-OCR-1994 (first entry)
XX MAB 022 VK chain.

FC receptor; FcR; humanized antibody; hAb; IgG; cancer; allergy;
autoimmune disease; heteroantibody; bifunctional antibody;
immunotoxin; CDR; complementarity determining region; VH;
heavy chain variable region; VK; kappa chain variable region;
mononuclear phagocyte; PCR; polymerase chain reaction; primer;
site-directed mutagenesis; monoclonal antibody; MAb.

Mus sp.
WO9410332-A.
11-MAY-1994.
04-NOV-1993; 93WO-US10384.
04-NOV-1992; 92GB-0023377.
(MEDA-) MEDAREX INC.

Carr FJ, Harris WJ, Tempest PR;
WPI; 1994-167486/20.

New humanised antibodies to Fc receptors - used for diagnosis or
for treatment of e.g. cancer, allergies and infectious and
auto-immune diseases

Disclosure; Page 23; 36pp; English.

Humanized antibodies (HABs) for IgG Fc receptors on human phagocytes
comprise the CDR of mouse monoclonal antibody 22 (from hybridoma
022MCL-1). VH chains from human Igs NEWM or KOL, and VK chains from
Ig REI. Sequences are provided for mouse 022 VH (AAR54931),
humanized NEWM-based VH (022 NMVH, AAR54929), humanized KOL-
based VH (022 KLVH, AAR54930), mouse 022 VK (AAR54933) and humanized
REI-based VK (022 HuVK, AAR54932). During hAb production, VH and VK
cDNAs were PCR amplified using primers given in AAO65378-87.
Mutagenesis of clone M13VHPCR2 KOL VH (L71R) was performed using
oligos AAO65388-89. The HABS can be used in heteroantibody,
bifunctional antibody and immunotoxin production.

SQ Sequence 112 AA;
Query Match 94.58; Score 551; DB 15; Length 112;
Best Local Similarity 93.84; Pred. No. 1.4e-39;
Matches 105; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
Qy 1 DIVVTQSPDSLAVSLGERVTMNCSSOSVLYSSNOKNYLAWYQKPGQSPKLLIYMASTR 60
Db 1 NIVVTQSPSSLAVSAGKRVKVTMSCKSSOSVLYSSNOKNYLAWYQKPGQSPKLLIYMASTR 60
Qy 61 ESGVPDRFSGSGGTFTLTISIVQAEDLAVYYQFYLSSYFGGGTKLEIK 112
Db 61 ESGVPDRFSGSGGTFTLTISIVQAEDLAVYYQFYLSSYFGGGTKLEIK 112
RESULT 2
ID AAW90226 standard; Protein: 263 AA.
XX AC AAW90226;
XX 10-MAY-1999 (first entry)
XX Anti-B7.2 monospecific triabody IG10.
KW B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
CD86; T cell activation; inhibitor; graft versus host disease;
transplant rejection; allograft rejection; autoimmune disease;
allergy; therapy; human; triabody; antibody; IG10.
XX Chimeric - Mus sp.
OS Chimeric - synthetic.
XX Key Location/Qualifiers
FT Peptide 1..24 /note= "pelB signal peptide"
FT Region 25..144 /note= "anti B7.2 MAB VH region"
FT Peptide 145..257 /note= "anti B7.2 MAB VL region"
FT Peptide 258..263 /note= "His6 tag"
XX WO9858965-A2.
XX 30-DEC-1998.
XX 22-JUN-1998; 98WO-EP03791.
XX 20-JUN-1997; 97EP-0870092.
XX (INNO-) INNOGENETICS NV.
XX Bosman A, Buysse M, Lorre K, Sablon E;
WPI; 1999-105615/09.
XX N-PSDB; AAX01660.
XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat
XX immune diseases including allograft rejection
XX Example 7.3; Fig 34; 182pp; English.
XX This polypeptide comprises a LG10 monospecific triabody composed
XX of the VH region of anti-B7.2 monoclonal antibody (MAB) IG10
XX joined to the VL region of IG10. A triabody is a mono- a bi- or
XX a trispecific molecule recognising simultaneously e.g. two B7.2
XX and one B7.1 molecules. It has a rigid structure that prevents
XX simultaneous binding to the 3 targets. Each antigen-binding site
XX is formed by pairing of one VH and one VL domain from the same or
XX from two different polypeptides. The invention relates to novel
XX molecules, including triabodies, which can cross-link and/or
XX cross-react with the costimulatory molecules B7.1 and B7.2 expressed

CC on professional antigen-presenting cells, leading to the inhibition
 CC of antigen-specific T cell activation. Methods are provided for
 CC the production of such B7-binding molecules, and for their use in
 CC the treatment or prevention of diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).

XX Sequence 263 AA;
 SQ Query Match 92.6%; Score 540; DB 20; Length 263;
 Best Local Similarity 92.0%; Pred. No. 2.8e-38;
 Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 DIVMTQSPDLSAVLSGERVTNCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR_60
 DB 145 DIETQSPSSLAVSAGEEVTNCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 204
 OY 61 ESVGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHOYLSSVTFGGGTGLEIK 112
 205 ESVGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHOYLSSVTFGGGTGLEIK 256

RESULT 3
 AAW90228
 ID AAW90228 standard; Protein; 268 AA.
 XX AC AAW90228;
 XX DT 10-MAY-1999 (first entry)
 XX DE Anti-B7.1/anti-B7.2 bispecific triabody II.
 XX KW B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 KW CD86; T cell activation; inhibitor; graft versus host disease;
 KW transplant rejection; allograft rejection; autoimmune disease;
 KW allergy; therapy; human; triabody; antibody; B7-24; IG10.
 XX OS Chimeric - Mus sp.
 OS Chimeric - synthetic.
 XX FH Key Location/Qualifiers
 FT Peptide 1..39
 FT /note= "g3p signal peptide"
 FT Region 40..155
 FT /note= "anti B7.1 Mab VH region"
 FT Peptide 156..268
 FT /note= "anti B7.2 Mab VL region"
 XX WO9858965-A2.
 PD 30-DEC-1998.
 XX PF 22-JUN-1998; 98WO-EP03791.
 XX PR 20-JUN-1997; 97EP-0870092.
 XX PA (INNO-) INNOGENETICS NV.
 XX PI Bosman A, Buyse M, Lorre K, Sablon E;
 XX WPI: 1999-105615/09.
 DR N-PSDB; AAX01662.
 XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 PT immune diseases including allograft rejection
 PS Example 7.3; Fig 38; 182pp; English.
 XX This polypeptide comprises a bispecific triabody composed of the VH
 CC region of anti-B7.1 monoclonal antibody (Mab) B7-24 joined to the
 CC VL region of anti-B7.2 Mab IG10. A triabody is a mono- a bi- or
 CC a trispecific molecule recognising simultaneously e.g. two B7.2
 CC and one B7.1 molecules. It has a rigid structure that prevents

CC simultaneous binding to the 3 targets. Each antigen-binding site
 CC is formed by pairing of one VH and one VL domain from the same or
 CC from two different polypeptides. The invention relates to novel
 CC molecules, including triabodies, which can cross-link and/or
 CC cross-react with the costimulatory molecules B7.1 and B7.2 expressed
 CC on professional antigen-presenting cells, leading to the inhibition
 CC of antigen-specific T cell activation. Methods are provided for
 CC the production of such B7-binding molecules, and for their use in
 CC the treatment or prevention of diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).

XX Sequence 268 AA;
 SQ Query Match 92.6%; Score 540; DB 20; Length 268;
 Best Local Similarity 92.0%; Pred. No. 2.9e-38;
 Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 DIVMTQSPDLSAVLSGERVTNCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
 DB 156 DIETQSPSSLAVSAGEEVTNCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 215
 OY 61 ESVGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHOYLSSVTFGGGTGLEIK 112
 DB 216 ESVGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHOYLSSVTFGGGTGLEIK 267

RESULT 4
 AAW90222
 ID AAW90222 standard; Protein; 268 AA.
 XX AC AAW90222;
 XX DT 10-MAY-1999 (first entry)
 XX DE Anti-B7.2 monospecific diabody IG-10.
 XX KW B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 KW CD86; T cell activation; inhibitor; graft versus host disease;
 KW transplant rejection; allograft rejection; autoimmune disease;
 KW allergy; therapy; human; diabody; antibody; IG-10.
 XX OS Chimeric - Mus sp.
 OS Chimeric - synthetic.
 XX FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /note= "peB signal peptide"
 FT Region 25..144
 FT /note= "anti B7.2 Mab VH region"
 FT Peptide 145..149
 FT /note= "G4S flexible linker"
 FT Region 150..262
 FT /note= "anti B7.2 Mab VL region"
 FT Peptide 263..268
 FT /note= "His6 tag"
 XX WO9858965-A2.
 PD 30-DEC-1998.
 XX PF 22-JUN-1998; 98WO-EP03791.
 XX PR 20-JUN-1997; 97EP-0870092.
 XX PA (INNO-) INNOGENETICS NV.
 XX PI Bosman A, Buyse M, Lorre K, Sablon E;
 XX WPI: 1999-105615/09.
 DR N-PSDB; AAX01656.
 XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat

PT Immune diseases including allograft rejection
 XX Example 7.2; Fig 26; 182pp; English.
 PS This polypeptide comprises a 10-10 monospecific diabody composed
 XX of the VH region of anti-B7.2 monoclonal antibody (MAB) 1G-10
 CC joined via a short, flexible linker to the VL region of 1G-10.
 CC Mono- or bispecific bivalent molecules are generated by shortening
 CC the flexible linker sequence between the VH and VL of the anti-B7.1
 CC scFv B7-24, the anti-B7.2 scFv 1G10 and the scFv molecule with
 CC dual specificity for B7.1 and B7.2 (B7.12) to 5-10 residues, and
 CC for bispecific molecules by cross-pairing the VH and VL domains
 CC from the 2 scFvs with different antigen recognition (B7.1/B7.2 and
 CC B7.12/B7.12). The invention relates to novel molecules, including
 CC diabolies, which can cross-link and/or cross-react with the
 CC costimulatory molecules B7.1 and B7.2 expressed on professional
 CC antigen-presenting cells, leading to the inhibition of antigen-
 CC specific T cell activation. Methods are provided for the
 CC production of such B7-binding molecules, and for their use in the
 CC treatment or prevention of diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (Claimed).
 XX
 SQ Sequence 268 AA;
 Query Match 92.6%; Score 540; DB 20; Length 268;
 Best Local Similarity 92.0%; Pred. No. 2.9e-38;
 Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DIVMTQSPDLSAVSLGERVTWCKSSQSVLYSSNOKNYLAWYQKPGSPKLLIYWASTR 60
 DB 150 DIETQSPSSSLAVSAGEEVTWCKSSQSVLYSSNOKNYLAWYQKPGSPKLLIYWASTR 209
 QY 61 ESGVPDRFGSGSGTDFLTITSSVQAEADLVVYCHQYLSSVTFGGTKLEIK 112
 DB 210 ESGVPDRFGSGSGTDFLTITSSVQAEADLVVYCHQYLSSVTFGGTKLEIK 261
 RESULT 5
 AAW90224
 ID AAW90224 standard; Protein; 273 AA.
 XX AAW90224;
 AC AAW90224;
 XX 10-MAY-1999 (first entry)
 DT Anti-B7.1/anti-B7.2 bispecific diabody II.
 DE B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 XX CD86; T cell activation; inhibitor; graft versus host disease;
 KW transplant rejection; allograft rejection; autoimmune disease;
 KW allergy; therapy; human; diabody; antibody; B7-24; 1G10.
 XX Chimeric - Mus sp.
 OS Chimeric - synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..39
 FT /note= "g3p signal peptide"
 FT Region 40..155
 FT /note= "anti B7.1 MAB VH region"
 FT Peptide 156..160
 FT /note= "G4S flexible linker"
 FT Region 161..273
 FT /note= "anti B7.2 MAB VL region"
 FT Misc-difference 21
 FT /note= "encoded by TCA"
 XX WO9858965-A2.
 XX 30-DEC-1998.
 XX 22-JUN-1998; 98WO-EP03791.
 XX

XX 20-JUN-1997; 97EP-0870092.
 XX (INNO-) INNOGENETICS NV.
 PA Bosman A, Buyse M, Lorre K, Sablon E;
 PI WPI: 1999-105615/09.
 XX N-PSDB; AAX01658.
 DR New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 PT immune diseases including allograft rejection
 XX Example 7.2; Fig 30; 182pp; English.
 CC This polypeptide comprises a bispecific diabody composed of the VH
 CC region of anti-B7.1 monoclonal antibody (MAB) B7-24 joined via a
 CC short, flexible linker to the VL region of anti-B7.2 MAB 1G10.
 CC Mono- or bispecific bivalent molecules are generated by shortening
 CC the flexible linker sequence between the VH and VL of the anti-B7.1
 CC scFv B7-24, the anti-B7.2 scFv 1G10 and the scFv molecule with
 CC dual specificity for B7.1 and B7.2 (B7.12) to 5-10 residues, and
 CC for bispecific molecules by cross-pairing the VH and VL domains
 CC from the 2 scFvs with different antigen recognition (B7.1/B7.2 and
 CC B7.12/B7.12). The invention relates to novel molecules, including
 CC diabolies, which can cross-link and/or cross-react with the
 CC costimulatory molecules B7.1 and B7.2 expressed on professional
 CC antigen-presenting cells, leading to the inhibition of antigen-
 CC specific T cell activation. Methods are provided for the
 CC production of such B7-binding molecules, and for their use in the
 CC treatment or prevention of diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (Claimed).
 XX
 SQ Sequence 273 AA;
 Query Match 92.6%; Score 540; DB 20; Length 273;
 Best Local Similarity 92.0%; Pred. No. 2.9e-38;
 Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DIVMTQSPDLSAVSLGERVTWCKSSQSVLYSSNOKNYLAWYQKPGSPKLLIYWASTR 60
 DB 161 DIETQSPSSSLAVSAGEEVTWCKSSQSVLYSSNOKNYLAWYQKPGSPKLLIYWASTR 220
 QY 61 ESGVPDRFGSGSGTDFLTITSSVQAEADLVVYCHQYLSSVTFGGTKLEIK 112
 DB 221 ESGVPDRFGSGSGTDFLTITSSVQAEADLVVYCHQYLSSVTFGGTKLEIK 272
 RESULT 6
 AAW90218
 ID AAW90218 standard; Protein; 556 AA.
 XX AAW90218;
 AC AAW90218;
 XX 10-MAY-1999 (first entry)
 DT Bispecific tetraivalent antibody BiTab1G10-B7-24H6.
 DE B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 XX CD86; T cell activation; inhibitor; graft versus host disease;
 KW transplant rejection; allograft rejection; autoimmune disease;
 KW allergy; therapy; human; bispecific tetraivalent antibody; BiTab;
 KW BiTab1G10-B7-24H6.
 XX Chimeric - Mus sp.
 OS Chimeric - Homo sapiens.
 OS Chimeric - synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..120
 FT /note= "VH region anti B7.2 MAB"
 FT Peptide 121..135

```

FT Region /note= "(G4S3) flexible linker"
FT 136..248
FT /note= "VL region anti B7.2 MAb"
FT 249..259
FT /note= "human IgG3 hinge region"
FT 260..285
FT /note= "helix-turn-helix dimerisation domain"
FT 286..305
FT /note= "human IgG3 hinge domain"
FT 306..426
FT /note= "VH region anti B7.1 MAB"
FT 427..441
FT /note= "(G4S3)flexible linker"
FT 442..550
FT /note= "VL region anti B7.1 MAB"
FT 551..556
FT /note= "His6 tag"
XX
XX WO9858965-A2.
PN
30-DEC-1998.
PF 22-JUN-1998; 98WO-EP03791.
XX
PR 20-JUN-1997; 97EP-0870092.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Bosman A, Buyse M, Lorre K, Sablon E;
XX WPI; 1999-105615/09.
DR N-PSDB; AAX01652.
XX
PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat
PT immune diseases including allograft rejection
XX
PS Example 7.1; Fig 18; 182pp; English.
XX
CC This polypeptide comprises the bispecific tetraivalent antibody
CC BiTab1G10-B7-24H6. The molecule consists of 4 scfvs, i.e. 2
CC anti B7.1 scfvs and 2 anti B7.2 scfvs (tetravalency). One single
CC BiTab is a homodimer of 2 identical molecules, each containing both
CC an anti B7.1 and anti B7.2 scfv (bispecificity). An anti-B7.1 and
CC anti-B7.2 scfv are linked using a dimerisation domain (see
CC AAW90219), which drives the homodimerisation of the molecule. DNA
CC (see AAX01652) encoding the BiTab has been constructed to allow
CC expression of the BiTab in transformed E. coli cells. The BiTab
CC cross-links, and/or cross-reacts, with the costimulatory molecules
CC B7.1 and B7.2 that are expressed on the membrane of professional
CC antigen-presenting cells, leading to the inhibition of antigen-
CC specific T cell activation. The invention relates to such
CC B7-binding molecules, methods for their production, and their use
CC for treating or preventing diseases of the immune system, in
CC particular graft rejection, graft versus host disease, allergy and
CC autoimmune diseases (claimed).
XX
SQ Sequence 556 AA:
Query Match 92.6%; Score 540; DB 20; Length 556;
Best Local Similarity 92.0%; Pred. No. 5.9e-38;
Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 DIVMTQSPDLSAVISGERVTMCKSSQSVLYSSNKNYLAHYQOKPGSPKLLIYWASTR 60
Db 136 DIELTQSPSSLAIVSAGEEVTMTCKSSQSVLYSSNKNYLAHYQOKPGSPKLLIYWASTR 195
QY 61 ESGVDPDRFSGSGGTFDFTLTISVQAEDLAVYYCHQYLSSYTFGGGTGLEIK 112
Db 196 ESGVDPDRFSGSGGTFDFTLTISVQAEDLAVYYCHQYLSSYTFGGGTGLEIK 247

```

RESULT 7
AAW90217

```

ID AAW90217 standard; Protein; 580 AA.
XX
AC AAW90217;
XX
DT 10-MAY-1999 (first entry)
XX
DE Bispecific tetraivalent antibody BiTabB7-24-Ig10H6.
XX
KW B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
KW CD86; T cell activation; inhibitor; graft versus host disease;
KW transplant rejection; allograft rejection; autoimmune disease;
KW allergy; therapy; human; bispecific tetraivalent antibody; BiTab;
KW BiTabB7-24-Ig10H6.
XX
OS Chimeric - Mus sp.
OS Chimeric - Homo sapiens.
OS Chimeric - Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT /note= "pe1B signal peptide"
FT Region 25..138
FT /note= "VH region anti B7.1 MAB"
FT Peptide 139..153
FT /note= "(G4S3) flexible linker"
FT Region 154..262
FT /note= "VL region anti B7.1 MAB"
FT Region 263..273
FT /note= "human IgG3 hinge region"
FT Domain 274..308
FT /note= "helix-turn-helix dimerisation domain"
FT Domain 309..319
FT /note= "human IgG3 hinge domain"
FT Region 320..446
FT /note= "VH region anti B7.2 MAB"
FT Peptide 447..461
FT /note= "(G4S3)flexible linker"
FT Region 462..574
FT /note= "VL region anti B7.2 MAB"
FT Peptide 575..580
FT /note= "His6 tag"
FT Misc-difference 261
FT /note= "encoded by CTC"
FT Misc-difference 322..327
FT /note= "codons for these amino acids are not
FT present in the DNA sequence for
FT BiTabB7-24-Ig1-H6 provided in the
FT specification"
XX
XX WO9858965-A2.
XX
PD 30-DEC-1998.
XX
PF 22-JUN-1998; 98WO-EP03791.
XX
PR 20-JUN-1997; 97EP-0870092.
XX
XX (INNO-) INNOGENETICS NV.
XX
PI Bosman A, Buyse M, Lorre K, Sablon E;
XX WPI; 1999-105615/09.
DR N-PSDB; AAX01651.
XX
PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat
PT immune diseases including allograft rejection
XX
PS Example 7.1; Fig 16; 182pp; English.
XX
CC This polypeptide comprises the bispecific tetraivalent antibody
CC BiTabB7-24-Ig10H6. The molecule consists of 4 scfvs, i.e. 2
CC anti B7.1 scfvs and 2 anti B7.2 scfvs (tetravalency). One single
CC BiTab is a homodimer of 2 identical molecules, each containing both

```

CC an anti B7.1 and anti B7.2 scFv (bispecificity). An anti-B7.1 and
 CC and anti-B7.2 scFv are linked using a dimerisation domain (see
 CC AAW90219), which drives the homodimerisation of the molecule. DNA
 CC (see AAX01651) encoding the B7.1 has been constructed to allow
 CC expression of the B7.1 in transformed E. coli cells. The B7.1
 CC cross-links, and/or cross-reacts, with the costimulatory molecules
 CC B7.1 and B7.2 that are expressed on the membrane of professional
 CC antigen-presenting cells, leading to the inhibition of antigen-
 CC specific T cell activation. The invention relates to such
 CC B7-binding molecules, methods for their production, and their use
 CC for treating or preventing diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).

XX Sequence 580 AA;
 SQ Query Match 92.6%; Score 540; DB 20; Length 580;
 Best Local Similarity 92.0%; Pred. No. 6.1e-38;
 Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 60
 II : IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
 Db 462 DIQLTQSPSSSLAVSAGEEVTMTCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 521
 61 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYCHQYLSSYTFGGTKLEIK 112
 IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
 Db 522 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYCHQYLSSYTFGGTKLEIK 573

RESULT 8
 AAR92215
 ID AAR92215 standard; Protein; 113 AA.
 XX AC AAR92215;
 XX DT 28-MAY-1996 (first entry)
 XX DE LL2 MAB VK region.
 XX KW Humanised antibody; monoclonal antibody; MAB; LL2; B-cell lymphoma;
 KW Leukaemia; therapy; diagnosis; complementarity determining region;
 KW CDR; antibody engineering.
 XX OS Mus musculus.
 XX FH Key
 FT Region 24..40 Location/Qualifiers
 /label= CDR1
 /note= "claim 6, page 44"
 56..62
 FT Region /label= CDR2
 /note= "claim 7, page 44"
 95..103
 FT Region /label= CDR3
 /note= "claim 8, page 44"
 FT FT
 XX WO9604925-A1.
 XX PN 22-FEB-1996.
 XX PD 11-AUG-1995; 95WO-US09641.
 XX PF 12-AUG-1994; 94US-0289576.
 XX PR (IMMU-) IMMUNOMEDICS INC.
 XX PA Hansen H, Leung S;
 XX PI WPI: 1996-139454/14.
 XX DR N-PSDB; AAT15802.
 XX PT Chimeric and humanised LL2 antibodies - used to produce conjugates
 for the therapy and diagnosis of B-cell lymphoma(s) and

PT Leukaemia(s).
 XX Claim 5; Page 35-36; 70pp; English.
 CC The complementarity determining regions (CDRs) of mouse monoclonal
 CC antibody (MAB) LL2 VK (AAR92215) and VH (AAR92216) regions were
 CC recombinantly linked to the framework sequences of human VK and VH
 CC regions, respectively, to give humanised LL2 VK (AAR92217) and VH
 CC (AAR92218). These were subsequently linked, respectively, to human
 CC kappa and IgG1 constant regions. A humanised MAB was obtained that
 CC retained the B-lymphoma and leukaemia cell targeting and
 CC internalisation characteristics of the parental LL2 MAB, and which
 CC exhibited a lowered HAMA reaction. It can be linked to e.g. a
 CC cytostatic agent for therapeutic appln.
 XX SQ Sequence 113 AA;
 Query Match 90.4%; Score 527; DB 17; Length 113;
 Best Local Similarity 88.4%; Pred. No. 1.6e-37;
 Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
 QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 60
 II : IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
 Db 1 DIQLTQSPSSSLAVSAGENVTMTCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 60
 61 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYCHQYLSSYTFGGTKLEIK 112
 IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
 Db 61 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYCHQYLSSYTFGGTKLEIK 112
 RESULT 9
 AAW27695
 ID AAW27695 standard; Protein; 113 AA.
 XX AC AAW27695;
 XX DT 14-APR-1998 (first entry)
 XX DE Variable kappa chain of MAB LL2.
 XX KW Variable kappa chain; B cell; monoclonal antibody; MAB; LL2;
 KW B cell lymphoma; lymphocytic leukaemia cell; murine;
 KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
 KW chronic lymphocytic leukaemia.
 XX OS Mus sp.
 XX OS Synthetic.
 XX FH Key
 FT Region 18..20 Location/Qualifiers
 /note= "potential N-linked glycosylation site"
 24..40
 FT Region /note= "complementarity determining region 1"
 56..62
 FT Region /note= "complementarity determining region 2"
 95..102
 FT Region /note= "complementarity determining region 3"
 XX WO9734632-A1.
 XX PN 25-SEP-1997.
 XX PD 19-MAR-1997; 97WO-US04196.
 XX PF 20-MAR-1996; 96US-0013709.
 XX PR (IMMU-) IMMUNOMEDICS INC.
 XX PA Hansen H, Leung S, Qu Z;
 XX PI WPI: 1997-479995/44.
 XX DR N-PSDB; AAT88128.
 XX PT

PT Monoclonal antibody engineered to contain glycosylation site - in
 PT non-Fc constant heavy or light chain region, useful to diagnose or
 PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
 XX
 PS Example 3; Fig 4A; 88pp; English.

XX The present sequence is the variable kappa chain of the
 CC B cell specific monoclonal antibody (Mab) LL2, which contains an
 CC engineered tri-peptide N-glycan acceptor sequence. LL2 is a highly
 CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
 CC murine Mab. The Mab can be used to diagnose or treat B
 CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
 CC lymphocytic leukaemia. The glycosylation site allows a label or
 CC therapeutic agent of increased size to be conjugated to the
 CC carbohydrate moiety, without affecting the Mab's binding affinity
 CC or specificity.

XX Sequence 113 AA;
 Query Match 90.4%; Score 527; DB 18; Length 113;
 Best Local Similarity 88.4%; Pred. No. 1.6e-37;
 Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
 DB 1 DIQLTQSPSLAVSAGENTMCKSSQSVLYSSANHKNYLAHYQKPGQSPKLLIYWASTR 60
 QY 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYCHQYLSSYTFGGGTKEIK 112
 DB 61 ESGVDPDRFSGSGGTDTLTITSRVQEDLAIYCHQYLSSWTFGGGTKEIK 112

RESULT 10
 AAB62769
 ID AAB62769 standard; Protein; 113 AA.
 AC AAB62769;
 XX
 DT 03-APR-2001 (first entry)
 XX Human HIV-1 monoclonal antibody SEQ ID NO: 68.
 DE
 XX Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;
 KW envelope glycoprotein; gp120; diagnosis.
 OS Homo sapiens.
 XX WO200100678-A1.

XX 04-JAN-2001.
 XX 23-JUN-2000; 2000WO-US17327.
 XX 30-JUN-1999; 99US-0141701.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Watkins BA, Reitz MS;
 XX WPI; 2001-112438/12.
 XX N-PSDB; AAF29070.
 XX Novel human monoclonal antibody immunoreactive with human
 PT Immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1
 PT in biological sample and providing passive immunotherapy to HIV-1
 PT Infected mammal

PS Claim 1; Page 65; 81pp; English.
 CC The present invention provides the protein and coding sequences for the
 CC variable regions of human monoclonal antibodies which are immunoreactive
 CC with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.
 CC These can be used in diagnosis and therapy of HIV-1 infection.

XX Sequence 113 AA;
 Query Match 90.2%; Score 526; DB 22; Length 113;
 Best Local Similarity 88.4%; Pred. No. 1.9e-37;
 Matches 99; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
 DB 1 DLVMTQSPDLSAVSLGERATINCKSSQSVLYSSNNKNYLAHYQKPGQSPKLLIYWASTR 60
 QY 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYCHQYLSSYTFGGGTKEIK 112
 DB 61 ESGVDPDRFSGSGGTDTLTITSSLOAEADVAVYCOQYFNTPTFGGTKEIK 112

RESULT 11
 AAR50313
 ID AAR50313 standard; Protein; 113 AA.
 AC AAR50313;
 XX
 DT 05-OCT-1994 (first entry)
 XX Humanised light chain variable region Pfhlzcl-1.
 DE Monoclonal antibody; Plasmodium falciparum; CDR;
 KW complementarity determining region; fusion protein;
 KW murine; variable; light; heavy; chain; malaria.
 XX Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 24..40
 FT /label= CDR1
 FT Region 56..62
 FT /label= CDR2
 FT Region 95..103
 FT /label= CDR3
 XX WO9405690-A.
 XX 17-MAR-1994.
 XX 08-SEP-1993; 93WO-US08435.
 XX 09-SEP-1992; 92US-0941654.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX (USNA) US SEC OF ARMY.
 XX (USNA) US SEC OF NAVY.
 XX Charoenvit Y, Hofman S, Hurle M, Rosenberg M;
 XX Sadooff JC, Sylvester DR, Gross MS;
 XX WPI; 1994-101115/12.
 XX N-PSDB; AAQ44827.
 XX New engineered antibodies and fusion proteins for preventing
 PT Plasmodium infection - contg. murine antibody CDR sequences, and
 PT corresp. nucleic acid, vectors and transformed cells
 XX Claim 5; Fig 2; 98pp; English.
 XX Naturally-occurring (AAQ44841-42) and synthetic (AAQ44825-28)
 CC variable light chain and variable heavy chain sequences
 CC derived from murine mAb NFS2 are provided. Murine mAb NFS2,
 CC its variable chain peptides, CDRs, functional fragments,
 CC Fab fragments, and analogs are useful in prodn. of fusion
 CC proteins, esp. engineered antibodies. These prods. are
 CC used to protect humans against Plasmodium infections.

XX Sequence 113 AA;

Query Match 90.1%; Score 525.5; DB 15; Length 113;
 Best Local Similarity 90.3%; Pred. No. 2.1e-37;
 Matches 102; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 1 DIVWTSQPSDLSVSLGRVTMCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60
 Db 1 DIVWTSQPSDLSVSLGRVTMCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60

Qy 61 ESGVPRFSGSGSGTDTLTITSSVQAEADLAVYYCHQVLS-SYTFGGGKLEIK 112
 Db 61 ESGVPRFSGSGSGTDTLTITSSVQAEADLAVYYCHQVLS-SYTFGGGKLEIK 113

RESULT 12
 AAG65565
 ID AAG65565 standard; protein; 120 AA;
 AC AAG65565;
 XX
 30-NOV-2001 (first entry)
 Amino acid sequence of protein seq Id No. 90.
 Gene library; Immunoglobulin; antibody library; human.
 Homo sapiens.
 OS
 XX
 PN W0200162907-A1.
 XX
 PD 30-AUG-2001.
 XX
 XX 22-FEB-2001; 2001WO-JP01298.
 XX
 XX 22-FEB-2000; 2000JP-0050543.
 PR
 XX (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.
 PA
 XX Kurosawa Y, Akahori Y, Iba Y, Morino K, Shinohara M, Takahashi M;
 PI Okuno Y, Shiraki K;
 PI
 XX WPI: 2001-565420/63.
 DR
 DR N-PSDB; AAH47729.
 XX
 XX Producing gene libraries and antibody libraries, involves selecting a
 PT light chain that binds to a heavy chain product to produce a functional
 PT formation, and producing a gene library of the light chain variable
 PT regions -
 XX
 XX Examples: p 165; 181pp; Japanese.
 CC The invention relates to producing gene libraries, comprising
 CC immunoglobulin light and heavy variable region. The method involves
 CC selecting light chain that binds with the heavy chain product to produce
 CC a functional conformation, producing a gene library comprising a
 CC collection of these light chain variable genes, and combining with gene
 CC library of heavy chain variable genes. The method is used for production
 CC of gene and antibody libraries.
 XX
 XX Sequence 120 AA;
 SQ

Query Match 90.0%; Score 524.5; DB 22; Length 120;
 Best Local Similarity 90.3%; Pred. No. 2.7e-37;
 Matches 102; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 1 DIVWTSQPSDLSVSLGRVTMCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60
 Db 1 DIVWTSQPSDLSVSLGRVTMCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60

Qy 61 ESGVPRFSGSGSGTDTLTITSSVQAEADLAVYYCHQVLS-YTFGGGKLEIK 112
 Db 61 ESGVPRFSGSGSGTDTLTITSSVQAEADLAVYYCHQVLS-YTFGGGKLEIK 113

RESULT 13
 AAR28807
 ID AAR28807 standard; Protein; 112 AA.
 XX
 XX AAR28807;
 XX
 DT 02-APR-1993 (first entry)
 XX
 DE 5A8 VL.
 XX
 XX Primer: polymerase chain reaction; amplifv: 5A8; heavy chain;
 KW variable; antibody: VH1BACK; VH1FOR; VH01; light chain; humanised;
 KW gp120; CD4; cell surface glycoprotein; CD4+; lymphocyte;
 KW HIV-induced syncytia formation.
 XX
 OS Synthetic.
 XX
 XX WO9209305-A.
 PN
 XX 11-JUN-1992.
 PD
 XX 27-NOV-1991; 91WO-US08843.
 PF
 XX 27-NOV-1990; 90US-0618542.
 PR
 XX (BIOJ) BIOGEN INC.
 PA
 XX Burkly LC, Chisholm PL, Rosa JJ, Rosa MD, Thomas DW;
 PI
 XX WPI: 1992-398399/48.
 DR
 DR N-PSDB; AAQ30881.
 XX
 XX New anti-CD4 antibody homologues - which bind CD4, do not block
 PT binding of HIV gp120 to CD4 but block HIV-induced syncytia
 PT formation between CD4+ cells
 PT
 XX Claim 27; Page 129-30; 205pp; English.
 PS
 XX The sequence represents the 5A8 light chain variable egion (5A8 VL)
 CC of the antibody of the invention. This sequence was inserted into
 CC the vector pUC19 aligned with the pMDR927 insert. The encoding
 CC DNA was amplified using the primer sequences given in AAQ30569-70. 5A8
 CC VL was used in conjunction with the 5A8 heavy chain variable region and
 CC these two chains were then humanised. The antibody produced was shown
 CC to bind to CD4 but did not block the binding of gp120 to CD4. CD4 is a
 CC cell surface glycoprotein of CD4+ lymphocytes (helper/inducer cells).
 CC The antibody blocked HIV-induced syncytia formation. This antibody
 CC can be used in the detection, prophylaxis and treatment of diseases
 CC caused by infective agents whose primary targets are CD4+ cells.
 XX
 SQ Sequence 112 AA;
 Query Match 89.4%; Score 521; DB 13; Length 112;
 Best Local Similarity 89.3%; Pred. No. 5e-37;
 Matches 100; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DIVWTSQPSDLSVSLGRVTMCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60
 Db 1 DIVWTSQPSDLSVSLGRVTMCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60

Qy 61 ESGVPRFSGSGSGTDTLTITSSVQAEADLAVYYCHQVLS-YTFGGGKLEIK 112
 Db 61 ESGVPRFSGSGSGTDTLTITSSVQAEADLAVYYCHQVLS-YTFGGGKLEIK 112

RESULT 14
 AAR50314
 ID AAR50314 standard; Protein; 113 AA.
 XX
 AC AAR50314;
 XX
 DT 05-OCT-1994 (first entry)

XX DE Humanised light chain variable region PfHzlcl²-2.
XX KW Monoclonal antibody; Plasmodium falciparum; CDR:
KW complementarity determining region; fusion protein;
KW murine; variable; light; heavy; chain; malaria.
XX OS
XX Synthetic.
XX FH Key Location/Qualifiers
XX FT Region 24..40
XX FT /label= CDR1
XX FT Region. 56..62
XX FT /label= CDR2
XX FT Region 95..103
XX FT /label= CDR3
XX PN WO9405690-A.
XX YY 17-MAR-1994.
XX PF 08-SEP-1993; 93WO-US08435.
XX XX 09-SEP-1992; 92US-0941654.
XX PR (SMIK) SMITHKLINE BEECHAM CORP.
XX PA (USSA) US SEC OF ARMY.
XX PA (USNA) US SEC OF NAVY.
XX XX
XX PI Charoenvit Y, Hoffman S, Hurlie M, Rosenberg M;
XX PI Sadoiff JC, Sylvestre DR, Gross MS;
XX DR WPI; 1994-101115/12.
XX DR N-PSDB; AAQ44828.
XX XX
XX PT New engineered antibodies and fusion proteins for preventing
XX PT Plasmodium infection - contg. murine antibody CDR sequences, and
XX PT corresp. nucleic acid, vectors and transformed cells
XX XX
XX PS Claim 5; Fig 3; 98pp; English.
XX PS
XX CC Naturally-occurring (AAQ4481-42) and synthetic (AAQ44825-28)
XX CC variable light chain and variable heavy chain sequences
XX CC derived from murine mAb NFS2 are provided. Murine mAb NFS2,
XX CC its variable chain peptides, CDRs, functional fragments,
XX CC Fab fragments, and analogs are useful in prodn. of fusion
XX CC proteins, esp. engineered antibodies. These prods. are
XX CC used to protect humans against Plasmodium infections.

Sequence 113 AA;

Query Match 89.3%; Score 520.5; DB 15; Length 113;
Best Local Similarity 89.4%; Pred. No. 5.5e-37;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQPSDLAVSLGERVTMCKSSQSVLYSSNKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQPSDLAVSLGERATINCKSSQSVLYSSNKNYLAWYQKPGQPPKLLIYWASTR 60
QY 61 ESGVPRFRSGSGSGDTFTLTISVQAEDLAVYYCHOYLS-SYTFGGGTTKLEIK 112
Db 61 ESGVPRFRSGSGSGDTFTLTISVQAEDLAVYYCQOQYSPRTFGGGTKVEIK 113

RESULT 15
AAW32483
ID AAW32483 standard; Protein; 155 AA.
XX AAW32483;
XX AC
XX 27-MAR-1998 (first entry)
XX DT
XX DE Kappa light chain variable region.

XX KW Catalytic antibody; growth factor; B-cell mitogenesis; LHL seq;
XX L protein; hen egg lysozyme; HEL; kappa light chain; LEN.
XX OS
XX Chimeric - Homo sapiens.
XX OS Chimeric - synthetic.
XX FH Key Location/Qualifiers
XX FT Peptide 22..29
XX FT /note= "FLAG epitope"
XX FT Peptide 147..155
XX FT /note= "strept-tag"
XX PN WO9735887-A1.
XX XX 02-OCT-1997.
XX XX 26-MAR-1997; 97WO-AU00194.
XX XX 27-FEB-1997; 97AU-0005375.
XX XX 26-MAR-1996; 96AU-0008951.
XX PA (AMRA-) AMRAD OPERATIONS PTY LTD.
XX XX
XX PI Koentgen F, Suess GM, Tarlington DM, Treutlein HR;
XX XX WPI; 1997-489572/45.
XX XX N-PSDB; AAW32483.
XX XX
XX PT New catalytic antibody precursors - comprising a B-cell surface
XX PT molecule binding portion which can induce B-cell mitogenesis
XX XX
XX PS Example 7; Page 73-74; 109pp; English.
XX XX
XX CC This polypeptide comprises the variable kappa light chain from
XX CC human myeloma protein LEN, flanked by an N-terminal FLAG epitope
XX CC (see AAW32484) and a C-terminal strep-tag (see AAW32485) to facilitate
XX CC purification from E. coli host cells. Kappa peptide is soluble at
XX CC relatively high concentrations and binds to protein L. Growth
XX CC factor TLHL (see AAW32482) was generated from LHL (see AAW32479), kappa
XX CC and linker sequences. It was designed so that the kappa portion of
XX CC the protein could be cleaved off by TEV protease, generating LHL.
XX CC Production of catalytic antibodies to a specific antigen comprises
XX CC administering to an animal a growth factor comprising an antigen
XX CC capable of interacting with a B cell bound catalytic antibody. The
XX CC antigen is fused to a B cell surface molecule binding protein for
XX CC the antigen to be cleaved and for the remainder of the molecule to
XX CC induce B cell mitogenesis (claimed). The catalytic antibodies
XX CC generated by the process can have diagnostic and therapeutic
XX CC applications.

Sequence 155 AA;

Query Match 89.3%; Score 520.5; DB 18; Length 155;
Best Local Similarity 89.4%; Pred. No. 7.5e-37;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQPSDLAVSLGERVTMCKSSQSVLYSSNKNYLAWYQKPGQSPKLLIYWASTR 60
Db 30 DIVMTQPSDLAVSLGERATINCKSSQSVLYSSNKNYLAWYQKPGQPPKLLIYWASTR 89
QY 61 ESGVPRFRSGSGSGDTFTLTISVQAEDLAVYYCHOYLS-SYTFGGGTTKLEIK 112
Db 90 ESGVPRFRSGSGSGDTFTLTISVQAEDLAVYYCQOQYSPYTFGGGTTKLEIK 142

RESULT 16
AAW06912
ID AAY06912 standard; Protein; 155 AA.
XX AAY06912;
XX AC
XX 01-JUL-1999 (first entry)
XX DT

XX DE Human variable kappa light chain protein.
 XX KW Growth factor precursor; B-cell surface; T cell surface; CAB; hepatitis;
 KW catalytic antibody; immunoglobulin; tumour necrosis factor; influenza;
 KW rheumatoid arthritis; toxic shock syndrome; viral docking receptor;
 KW human immune virus; tumour-specific antigen; amyloid plaque; myeloma;
 KW Alzheimer's disease; IGE; allergy; asthma; drug detoxification; LEN;
 KW autoimmune; inflammatory disease; gene therapy; human; myeloma protein;
 KW kappa light chain protein.
 XX OS Homo sapiens.
 XX OS WO9915563-A1.
 XX PN 01-APR-1999.
 XX PD 18-SEP-1998; 98WO-AU00783.
 XX PF 19-SEP-1997; 97AU-0009306.
 XX PR (AMRA-) AMRAD OPERATIONS PTY LTD.
 XX PI Koentgen F, Suess GM, Tarlington DM, Treutlein HR;
 XX DR WPI; 1999-244394/20.
 XX DR N-PSDB; AAX34594.
 XX PT Growth factor precursor cleaved by antigen-specific catalytic
 XX PT antibody
 XX PS Example 7; Page 72-73; 101pp; English.

XX CC The invention relates to a growth factor precursor that comprises B-cell
 CC surface binding part, T cell surface binding part, antigen cleavable
 CC by a catalytic antibody (CAB); and a peptide comprising heavy and light
 CC chains of immunoglobulin. When the antigen is cleaved the B cell surface
 CC part can interact with its target. The growth factor precursors are used
 CC to select B cells that produce Ag-specific CAB, and to generate CAB from
 CC such cells (by inducing mitogenesis, caused by the growth factor released
 CC by specific cleavage). The Ag-specific CAB can be directed against, e.g.
 CC tumour necrosis factor (treatment of rheumatoid arthritis or toxic shock
 CC syndrome); viral docking receptors (treatment of human immune virus,
 CC hepatitis and influenza infections); tumour-specific antigens; amyloid
 CC plaque (treatment of Alzheimer's disease or myeloma) or IGE (treatment of
 CC allergies such as asthma). CAB may also be used for drug detoxification,
 CC to treat autoimmune or inflammatory diseases and to eliminate
 CC environmental or industrial pollutants, such as plastics and petroleum.
 CC Particularly the growth factor precursors are produced by delivering the
 CC corresponding nucleic acid in a viral or other gene therapy vector. The
 CC present sequence represents the amino acid sequence of the human variable
 CC kappa light chain. This was generated based on the human myeloma protein
 CC LEN sequence.

XX SQ Sequence 155 AA;
 Query Match 89.3%; Score 520.5; DB 20; Length 155;
 Best Local Similarity 89.4%; Pred. No. 7.5e-37;
 Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYOQKPCQSPKLLIYWASTR 60
 DB 30 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNKNYLAAYOQKPCQSPKLLIYWASTR 89
 QY 61 ESGVPDRFSGSGGTDTFTLTSSVQAEADLAVYCHQYLS- YTFGGGKLEIK 112
 DB 90 ESGVPDRFSGSGGTDTFTLTSSVQAEADLAVYCHQYLS- YTFGGGKLEIK 142

RESULT 17
 AAW32482
 ID AAW32482 standard; Protein; 342 AA.

AC AAW32482;
 XX 27-MAR-1998 (first entry)
 XX DE Growth factor TLHL (catalytic antibody precursor).
 XX DE Catalytic antibody; growth factor; B-cell mitogenesis; TLHL;
 KW L protein; hen egg lysozyme; kappa light chain; ss.
 XX OS Chimeric - Peptostreptococcus magnus.
 OS Chimeric - Gallus sp.
 OS Chimeric - Homo sapiens.
 XX PN WO9735887-A1.
 XX PD 02-OCT-1997.
 XX PF 26-MAR-1997; 97WO-AU00194.
 XX PR 27-FEB-1997; 97AU-0005375.
 XX PR 26-MAR-1996; 96AU-0008951.
 XX PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 XX PI Koentgen F, Suess GM, Tarlington DM, Treutlein HR;
 XX DR WPI; 1997-489572/45.
 XX DR N-PSDB; AAT91589.
 XX PT New catalytic antibody precursors - comprising a B-cell surface
 XX PT molecule binding portion which can induce B-cell mitogenesis
 XX PS Example 11; Page 66-68; 109pp; English.

XX CC This protein comprises growth factor TLHL, where L is the
 CC immunoglobulin binding entity from Peptostreptococcus magnus, H is
 CC residues 42-62 of hen egg lysozyme, and T represents the variable
 CC kappa light chain (see AAW32483) from human myeloma protein LEN. It
 CC was expressed in E. coli DH10B cells utilising a DNA construct (see
 CC AAT91589) produced from LHL (see AAT91986), kappa (see AAT91590) and
 CC synthetic linker oligonucleotides, and was purified from total
 CC bacterial lysate via a strep-tag. The linker contains a cleavage
 CC site for tobacco etch virus (TEV) protease. TLHL was designed so
 CC that the kappa portion of the protein is cleaved by TEV protease
 CC into 'T' and 'LH' moieties. Production of catalytic antibodies to
 CC a specific antigen comprises administering to an animal a growth
 CC factor comprising an antigen capable of interacting with a B cell
 CC bound catalytic antibody. The antigen is fused to a B cell surface
 CC molecule binding protein for the antigen to be cleaved and for the
 CC remainder of the molecule to induce B cell mitogenesis (claimed).
 CC LHL crosslinks with surface immunoglobulin on B cells. This induces
 CC B cell activation and blast formation. Internalisation and
 CC processing of LHL leads to the presentation of the H peptide on
 CC MHC II. T cell recognition of MHC II with the H peptide signals the
 CC activated B cell to proliferate and undergo antibody class switching
 CC and secretion. The catalytic antibodies generated by the process
 CC can have diagnostic and therapeutic applications.

XX SQ Sequence 342 AA;
 Query Match 89.3%; Score 520.5; DB 18; Length 342;
 Best Local Similarity 89.4%; Pred. No. 1.6e-36;
 Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYOQKPCQSPKLLIYWASTR 60
 DB 30 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNKNYLAAYOQKPCQSPKLLIYWASTR 89
 QY 61 ESGVPDRFSGSGGTDTFTLTSSVQAEADLAVYCHQYLS- YTFGGGKLEIK 112
 DB 90 ESGVPDRFSGSGGTDTFTLTSSVQAEADLAVYCHQYLS- YTFGGGKLEIK 142

RESULT 18
AA06909
ID AAY06909 standard; Protein; 342 AA.
XX
AC AAY06909;
XX
DT 01-JUL-1999 (first entry)
XX
DE TLHL amino acid sequence.
XX
KW Growth factor precursor; B-cell surface; T cell surface; CAB; hepatitis;
KW catalytic antibody; immunoglobulin; tumour necrosis factor; influenza;
KW rheumatoid arthritis; toxic shock syndrome; viral docking receptor;
KW human immune virus; tumour-specific antigen; amyloid plaque; myeloma;
KW Alzheimer's disease; IGE; allergy; asthma; drug detoxification;
KW autoimmune; inflammatory disease; gene therapy; protein L; P. magnus;
KW hen egg lysozyme; HEL; LHL; TLHL.
XX
DE Peptostreptococcus magnus.
XX
ID AAW32480
XX
AC AAW32480 standard; Protein; 495 AA.
XX
AC AAW32480;
XX
DT 27-MAR-1998 (first entry)
XX
DE Growth factor CATAB-TEV (catalytic antibody precursor).
XX
KW Catalytic antibody; growth factor; B-cell mitogenesis; LHL;
KW L protein; hen egg lysozyme; kappa light chain.
XX
OS Chimeric - Peptostreptococcus magnus.
OS Chimeric - Gallus sp.
OS Chimeric - Homo sapiens.
XX
PN WO9735887-A1.
XX
PD 02-OCT-1997.
XX
PF 26-MAR-1997; 97WO-AU00194.
XX
PR 27-FEB-1997; 97AU-0003375.
PR 26-MAR-1996; 96AU-0008951.
XX
PA (AMRA-) AMRAD OPERATIONS PTY LTD.
XX
PI Koentgen F, Sues GM, Tarlington DM, Treutlein HR;
DR WPI: 1999-244394/20.
DR N-PSDB; AAX34592.
XX
PT Growth factor precursor cleaved by antigen-specific catalytic
PT antibody
XX
XX Example 11; Page 67-69; 101pp; English.
XX
CC The invention relates to a growth factor precursor that comprises B-cell
CC surface binding part, T cell surface binding part, antigen cleavable
CC by a catalytic antibody (CAB); and a peptide comprising heavy and light
CC chains of immunoglobulin. When the antigen is cleaved the B cell surface
CC part can interact with its target. The growth factor precursors are used
CC to select B cells that produce Ag-specific CAB, and to generate CAB from
CC such cells (by inducing mitogenesis, caused by the growth factor released
CC by specific cleavage). The Ag-specific CAB can be directed against, e.g.
CC tumour necrosis factor (treatment of rheumatoid arthritis or toxic shock
CC syndrome); viral docking receptors (treatment of human immune virus,
CC hepatitis and influenza infections); tumour-specific antigens; amyloid
CC plaque (treatment of Alzheimer's disease or myeloma) or IGE (treatment of
CC allergies such as asthma). CAB may also be used for drug detoxification,
CC to treat autoimmune or inflammatory diseases and to eliminate
CC environmental or industrial pollutants, such as plastics and petroleum.
CC Particularly the growth factor precursors are produced by delivering the
CC corresponding nucleic acid in a viral or other gene therapy vector. The
CC present sequence represents the amino acid sequence of TLHL. The LHL is a
CC growth factor comprising H flanked by two L molecules where L is a B cell
CC surface molecule binding portion from protein L of P. magnus; H is a T
CC cell surface molecule binding portion from hen egg lysozyme (HEL). TLHL
CC is LHL linked to kappa chain via TEV (tobacco etch virus) sensitive
CC peptide and g attached to N terminus region.
XX
SQ Sequence 342 AA;
Query Match 89.3%; Score 520.5; DB 20; Length 342;
Best Local Similarity 89.4%; Pred. No. 1.6e-36;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
1 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAWYQQKPGQPKLLIYWASTR 60
|||||
30 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAWYQQKPGQPKLLIYWASTR 89
61 ESGVDPDRFSGSGTDFLTITSSVQAEADLAVVYCHQYLSS-YTFGGGTGLEIK 112
|||||
90 ESGVDPDRFSGSGTDFLTITSSVQAEADLAVVYCHQYLSS-YTFGGGTGLEIK 142
|||||
RESULT 19
AAW32480
ID AAW32480 standard; Protein; 495 AA.
XX
AC AAW32480;
XX
DT 27-MAR-1998 (first entry)
XX
DE Growth factor CATAB-TEV (catalytic antibody precursor).
XX
KW Catalytic antibody; growth factor; B-cell mitogenesis; LHL;
KW L protein; hen egg lysozyme; kappa light chain.
XX
OS Chimeric - Peptostreptococcus magnus.
OS Chimeric - Gallus sp.
OS Chimeric - Homo sapiens.
XX
PN WO9735887-A1.
XX
PD 02-OCT-1997.
XX
PF 26-MAR-1997; 97WO-AU00194.
XX
PR 27-FEB-1997; 97AU-0003375.
PR 26-MAR-1996; 96AU-0008951.
XX
PA (AMRA-) AMRAD OPERATIONS PTY LTD.
XX
PI Koentgen F, Sues GM, Tarlington DM, Treutlein HR;
DR WPI: 1997-489572/45.
DR N-PSDB; AAT91587.
XX
PT New catalytic antibody precursors - comprising a B-cell surface
PT molecule binding portion which can induce B-cell mitogenesis
XX
XX Claim 12; Page 61-63; 109pp; English.
XX
CC This polypeptide comprises CATAB-TEV, a synthetic growth factor
CC composed of: (a) TLHL (see AAW91589), where L is the immunoglobulin
CC binding entity of Peptostreptococcus magnus, H is residues 42-62 of
CC hen egg lysozyme, and T represents the variable kappa light chain
CC (see AAW32483) from human myeloma protein LEN; and (b) an additional,
CC C-terminal kappa protein, the elements of CATAB-TEV being joined by
CC linkers containing recognition sites for tobacco etch virus (TEV).
CC protease. CATAB-TEV was designed so that the kappa portions of the
CC protein are removed by TEV protease to release 'T' and 'LHL'.
CC moieties. Production of catalytic antibodies to a specific antigen
CC comprises administering to an animal a growth factor comprising an
CC antigen capable of interacting with a B cell bound catalytic
CC antibody. The antigen is fused to a B cell surface molecule
CC binding protein for the antigen to be cleaved and for the remainder
CC of the molecule to induce B cell mitogenesis (claimed). LHL (see
CC AAW32479) crosslinks with surface immunoglobulin on B cells. This
CC induces B cell activation and blast formation. Internalisation
CC and processing of LHL leads to the presentation of the H peptide on
CC MHC II. T cell recognition of MHC II with the H peptide signals the
CC activated B cell to proliferate and undergo antibody class switching
CC and secretion. The catalytic antibodies generated by the process
CC can have diagnostic and therapeutic applications.
XX
SQ Sequence 495 AA;
Query Match 89.3%; Score 520.5; DB 18; Length 495;
Best Local Similarity 89.4%; Pred. No. 2.4e-36;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVVTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAUWQKPGQSPKLLIYWASTR 60
 Db 30 DIVVTQSPDSLAVSLGERATINCKSSQSVLYSSNQKNYLAUWQKPGQSPKLLIYWASTR 89
 QY 61 ESGVPDRFSGSGGTDFTLTSSVQAEADLAVIYCHQYLSS-YTFGGGKLEIK 112
 Db 90 ESGVPDRFSGSGGTDFTLTSSVQAEADLAVIYCHQYLSS-YTFGGGKLEIK 142

RESULT 20
 AAY06908
 ID AAY06908 standard; Protein; 495 AA.

XX AC AAY06908;
 XX DT 01-JUL-1999 (first entry)
 XX DE CATAB-TEV aminoacid sequence.

XX Growth factor precursor; B-cell surface; T cell surface; CAB; hepatitis;
 catalytic antibody; immunoglobulin; tumour necrosis factor; influenza;
 rheumatoid arthritis; toxic shock syndrome; viral docking receptor;
 human immune virus; tumour-specific antigen; amyloid plaque; myeloma;
 Alzheimer's disease; IGE; allergy; asthma; drug detoxification;
 autoimmune; inflammatory disease; gene therapy; protein L; P. magnus;
 hen egg lysosyme; HEL; LHL; CATAB-TEV; tobacco etch virus; TLHL.

XX Peptostreptococcus magnus.
 OS Gallus sp.

XX W09915563-A1.

XX 01-APR-1999.

XX 18-SEP-1998; 98WO-AU00783.

XX 19-SEP-1997; 97AU-0009306.

XX (AMRA-) AMRAD OPERATIONS PTY LTD.

XX Koentgen F, Suess GM, Tarlington DM, Treutlein HR;

XX WPI; 1999-244394/20.

XX N-PSDB; AAX34591.

XX Growth factor precursor cleaved by antigen-specific catalytic
 antibody

XX Example 15; Page 64-66; 101pp; English.

CC The invention relates to a growth factor precursor that comprises B-cell
 surface binding part, T cell surface binding part, antigen cleavable
 by a catalytic antibody (CAB); and a peptide comprising heavy and light
 chains of immunoglobulin. When the antigen is cleaved the B cell surface
 part can interact with its target. The growth factor precursors are used
 to select B cells that produce Ag-specific CAB, and to generate CAB from
 such cells (by inducing mitogenesis, caused by the growth factor released
 by specific cleavage). The Ag-specific CAB can be directed against, e.g.
 tumour necrosis factor (treatment of rheumatoid arthritis or toxic shock
 syndrome); viral docking receptors (treatment of human immune virus,
 hepatitis and influenza infections); tumour-specific antigens; amyloid
 plaque (treatment of Alzheimer's disease or myeloma) or IGE (treatment of
 allergies such as asthma). CAB may also be used for drug detoxification,
 to treat autoimmune or inflammatory diseases and to eliminate
 environmental or industrial pollutants, such as plastics and petroleum.
 CC Particularly the growth factor precursors are produced by delivering the
 corresponding nucleic acid in a viral or other gene therapy vector. The
 present sequence represents the amino acid sequence of CATAB-TEV which
 is assembled from TLHL and kappa by PCR. The LHL is a growth factor
 comprising H flanked by two L molecules where L is a B cell surface
 molecule binding portion from protein L of P. magnus; H is a T cell
 surface molecule binding portion from hen egg lysosyme (HEL). TLHL is

CC LHL linked to kappa chain via TEV (tobacco etch virus) sensitive
 CC peptide and g attached to N terminus region.

XX Sequence 495 AA;

XX Query Match 89.3%; Score 520.5; DB 20; Length 495;
 Best Local Similarity 89.4%; Pred. No. 2.4e-36;
 Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVVTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAUWQKPGQSPKLLIYWASTR 60
 Db 30 DIVVTQSPDSLAVSLGERATINCKSSQSVLYSSNQKNYLAUWQKPGQSPKLLIYWASTR 89
 QY 61 ESGVPDRFSGSGGTDFTLTSSVQAEADLAVIYCHQYLSS-YTFGGGKLEIK 112
 Db 90 ESGVPDRFSGSGGTDFTLTSSVQAEADLAVIYCHQYLSS-YTFGGGKLEIK 142

RESULT 21

AAY95243
 ID AAY95243 standard; Protein; 137 AA.

XX AC AAY95243;

XX 29-AUG-2000 (first entry)

XX Humanised antibody HuCC49 light chain variable region.

XX Humanised antibody; monoclonal antibody; CC49; HuCC49; CDR;
 complementarity determining region; mouse; human; carcinoma;
 colon cancer; tumor associated glycoprotein-72; TAG-72;
 tumour marker; diagnosis; therapy.

XX Chimeric - Mus musculus.

XX Chimeric - Homo sapiens.

XX Key Location/Qualifiers

FT Region 44..59 /note= "CDR1"

FT Region 76..82 /note= "CDR2"

FT Region 115..123 /note= "CDR3"

XX W0200026394-A1.

XX 11-MAY-2000.

XX 29-OCT-1999; 99WO-US25552.

XX 31-OCT-1998; 98US-0106534.

XX 02-NOV-1998; 98US-0106757.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Kashmiri SVS, Padlan EA, Schlom J;

XX WPI; 2000-365637/31.

XX Chimeric variants of CC49 monoclonal antibodies useful for detecting
 and treating cancers associated with the expression of the pancarcinoma
 tumor-associated antigen TAG-72 -
 Disclosure; Fig 4; 76pp; English.

XX The present sequence is that of the light chain variable region
 (VL) of huCC49, a humanised monoclonal antibody (MAB) formed by
 CC grafting hypervariable regions from murine MAB CC49 into VL and VH
 CC frameworks of human Mabs LEN and 21/28' CL, respectively, while
 CC retaining murine framework residues required for integrity of the
 CC antigen combining site structure. HuCC49 binds to the human
 CC pancarcinoma tumor associated glycoprotein-72 (TAG-72), which is
 CC found on the surface of certain human tumours. The invention is

PT	binding of HIV gp120 to CD4 but block HIV-induced syncytia formation between CD4+ cells
XX	
PS	Disclosure; Page 166-7; 205pp; English.
XX	
CC	The sequence given is encoded by the insert of the vector pMDR1007.
CC	pMDR1006 (see AAQ30919) and pSAB132 (see AAQ30906) were used in the construction of this vector. Three fragments were ligated together
CC	to generate pMDR1006; a 572 bp fragment of pMDR985 (see AAQ30913), a
CC	344bp AatII/EcoRV fragment of pMDR986 (see AAQ30918) and a 326 bp
CC	EcoRV/BglIII fragment of pMDR1003 (see AAQ30900). The ligation mixture
CC	was used to transform E. coli JA221(Iq) to ampicillin resistance.
CC	pMDR1006 was cleaved with NotI and the 1693 bp fragment isolated was
CC	ligated into NotI linearised pSAB132 which had been dephosphorylated
CC	by calf alkaline phosphatase. This generates the plasmid pMDR1007.
CC	The dephosphorylated mixture was fractionated through low temperature
CC	melting agarose and used to transform E. coli JA221(Iq) to ampicillin
CC	resistance. The pMDR1007 insert comprises DNA encoding, in a 5' to 3'
CC	order, the immunoglobulin kappa chain signal peptide, amino acid (AA)
CC	1-AA112 of the humanised 5A8 light chain variable region (LV) followed
CC	by genomic DNA encoding AA108-AA214 of the human kappa light chain,
CC	i.e. the light chain constant region (LC). This polypeptide is an
CC	antibody homolog which was shown to bind to CD4 but did not block the
CC	binding of gp120 to CD4. CD4 is a cell surface glycoprotein of CD4+
CC	lymphocytes (helper/inducer cells). The homolog blocked HIV-induced
CC	syncytia formation. This homolog can be used in the detection,
CC	prophylaxis and treatment of diseases caused by infective agents whose
CC	primary targets are CD4+ cells.
XX	
SQ	Sequence 241 AA;
	Query Match 88.3%; Score 515; DB 13; Length 241;
	Best Local Similarity 88.4%; Pred No. 3.4e-36;
	Matches 99; Conservative 5; Mismatches 8; Indels 0; Gaps
QY	1 DIVMTQSPDLSAVISGERVNTNCKSSQSVLYSSNOKNYLAWYQQKPGQSPKLLITYWASTR
DB	23 DIVMTQSPDLSAVISGERATINCKSSGLLYSTNKNYLAWYQQKPGQPPELLIYWASTR
QY	61 ESGVPDRFSGSGGDFTLTITSSVOAEDLAVYYCHQVLSSYTGGGTGLEIK 112
DB	83 ESGVPDRFSGSGGDFTLTITSSQAEDVAVYYCQYYRYFTFGRTGLEIK 134
RESULT 23	
AAB09779	
ID	AAB09779 standard; Protein; 259 AA.
XX	AAB09779;
AC	
XX	
DT	06-SEP-2000 (first entry)
XX	
DE	Antiviral scFv-antibody against a plant virus minimal protein.
XX	
KW	Molecular pathogenicide; plant disease; resistance; antibody; scFv;
KW	gene construct; pathogen; toxin; fusion protein; antimicrobial;
KW	deoxyribonuclease; RNase; ribosome inactivator; immunomodulator.
OS	Tobacco mosaic virus.
OS	Synthetic.
PN	WO200023593-A2.
XX	
PD	27-APR-2000.
PF	
XX	15-OCT-1999; 99WO-EP07844.
PR	
XX	16-OCT-1998; 98EP-0119630.
PR	16-OCT-1998; 98IN-0000666.
XX	(FRAU) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.
XX	Fischer R., Schillberg S., Naehring J., Sack M., Moncke M., Liao Y.

PI Spiegel H, Zimmerman S, Emans N, Holzem A;
XX WPI; 2000-339692/29.
XX
XX New fusion proteins and gene constructs for expressing agents
PT (antibodies, enzymes, vectors or molecular pathogenecides), useful for
PT protecting plants against pathogens and increasing resistance to
PT disease
XX
XX
PS Example 6; Fig 23c; 193pp; English.
XX
XX The present invention describes a fusion protein (I) comprising at least
CC one binding domain specifically recognising an epitope of a plant
CC pathogen and at least one further domain comprising a protein or peptide
CC sequence which is toxic to the pathogen or detrimental to its
CC replication, transmission or life cycle. Also described is a
CC pathogenecide (II) comprising (i) and a cellular targeting sequence
CC and/or membrane localisation sequence and/or motif that leads to
CC membrane anchoring; or at least one binding domain that specifically
CC recognises a viral movement and/or replicase protein. The fusion
CC protein, pathogenecide, polynucleotide, vectors, and compositions from
CC the present invention are useful for the protection of a plant against
CC the action of a pathogen. The kit from the present invention is useful
CC for carrying out the methods and may be employed in different
CC applications, for example in the diagnostic field or as research tools.
CC The kit or its components, such as the fusion protein, pathogenecide,
CC polynucleotides, vectors or compositions are useful in plant cell and
CC plant tissue culture, in agriculture. They are extremely useful for
CC breeding new varieties of plants that display improved properties such as
CC resistance to pathogens. AA56587 to AA56702 and AAB09774 to B097820
CC represent sequences used in the exemplification of the present
CC invention.
XX
XX Sequence 259 AA;
SQ
Query Match 88.0%; Score 513; DB 21; Length 259;
Best Local Similarity 86.6%; Pred. No. 5.4e-36;
Matches 97; Conservative 9; Mismatches 6; Indels 0; Gaps 0;
QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Db 141 DIVLSQSPDLSAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 200
QY 61 ESGVDPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLSSYTFGGGTGLEIK 112
Db 201 ESGVDPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLSSYTFGGGTGLEIK 252
LT 24
0894
AAP80894 standard; protein; 133 AA.
XX
XX AAP80894;
XX
XX 03-DEC-1990 (first entry)
XX
XX V region of L chain of anti-P.aeruginosa exotoxin Ab #1.
XX
XX Pseudomonas aeruginosa; anti-exotoxin antibody; L chain; V region;
KW ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..20
FT /label-signal peptide
FT 21..133
FT /label-V region of L chain
XX
XX EP270077-A.
XX
XX 08-JUN-1988.
PD
XX

PF 01-DEC-1987; 87EP-0117760.
XX
XX 03-DEC-1986; 86JP-0288340.
PR 26-NOV-1987; 87JP-0298513.
XX
XX (SUMO) SUMITOMO CHEM IND KK.
XX
XX Nakatani T, Nomura N, Horigome K, Noguchi H;
XX
XX WPI; 1988-156310/23.
DR N-PSDB; AAN80499.
XX
XX New gene encoding for antibody to Pseudomonas aeruginosa exotoxin
PT plus recombinant vectors and host cells, useful for treating
PT infections.
XX
XX Claim 4; Page 25; 39pp; English.
PS
XX Sequence is variable region of light chain of anti-exotoxin
CC antibody with signal sequence. AAN80498 encodes the same sequence
CC except that its signal peptide-encoding sequence contains an
CC intron.
CC See also AAN80495-N80496, AAN80498 and AAN80941-2.
XX
XX Sequence 133 AA;
SQ
Query Match 87.9%; Score 512.5; DB 9; Length 133;
Best Local Similarity 88.5%; Pred. No. 3.1e-36;
Matches 100; Conservative 5; Mismatches 7; Indels 1; Gaps 1;
QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Db 21 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNLYLAWYQKPGQSPKLLIYWASTR 80
QY 61 ESGVDPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLSS-YTFGGGTGLEIK 112
Db 81 ESGVDPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLSS-YTFGGGTGLEIK 133
RESULT 25
AAY50690
ID AAY50690 standard; Protein; 134 AA.
XX
XX AAY50690;
XX
XX 09-FEB-2000 (first entry)
XX
XX Human Hum4 VL ClaI-HindIII segment encoded protein.
XX
XX Human; antibody; humanized; anti-tumor; sialylated glycoprotein antibody;
KW TAG-72; cytostatic; cancer antigen; detection; carcinoma lesion;
KW diagnostic; treatment.
XX
XX Homo sapiens.
XX
XX US5976531-A.
XX
XX 02-NOV-1999.
XX
XX 16-JUN-1994; 94US-0261354.
XX
XX 19-APR-1990; 90US-0510697.
PR 20-OCT-1992; 92US-0964536.
XX
XX (DOWC) DOW CHEM CO.
PA
XX Johnson KS, Mezes PS, Richard RA;
XX
XX WPI; 1999-632731/54.
DR N-PSDB; AAZ23970.
XX
XX New humanized anti-TAG-72 antibodies, used for the detection, in vivo
PT imaging and treatment of cancers

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:00:06 ; Search time 6.55319 Seconds
(without alignments)
502.864 Million cell updates/sec

Title: US-10-056-052A-l8

Perfect score: 583

Sequence: 1 DIVMTQSPDLSAVSLGERVT.....YCHOYLSSTFGGTTKLEIK I12

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	527	90.4	113	1	US-08-690-102A-2
2	527	90.4	113	4	US-09-127-902-2
3	527	90.4	113	5	PCT-US95-09641-2
4	525	90.1	113	5	PCT-US93-08435-6
5	521	89.4	112	1	US-07-942-245-30
6	521	89.4	112	2	US-07-916-098A-15
7	520.5	89.3	113	5	PCT-US93-08435-8
8	520.5	89.3	113	3	US-08-828-741B-11
9	520.5	89.3	113	3	US-09-160-567-11
10	520.5	89.3	113	3	US-08-828-741B-6
11	520.5	89.3	113	3	US-09-160-567-6
12	520.5	89.3	113	3	US-08-828-741B-4
13	520.5	89.3	113	3	US-08-828-741B-4
14	520.5	89.3	113	3	US-08-828-741B-4
15	515	88.3	241	2	US-07-916-098A-56
16	513.5	88.1	113	4	US-08-525-539A-80
17	512.5	87.9	113	3	US-08-463-903-4
18	512.5	87.9	113	3	US-07-935-695-4
19	512.5	87.9	113	3	US-08-463-903-6
20	512.5	87.9	113	3	US-07-935-695-6
21	511.5	87.7	114	4	US-09-025-769B-17
22	510.5	87.6	113	3	US-08-483-749A-16
23	510.5	87.6	171	3	US-08-463-903-20
24	510.5	87.6	171	4	US-07-935-695-20
25	507.5	87.0	114	1	US-08-360-125-6
26	507.5	87.0	114	2	US-08-450-578-6
27	507.5	87.0	114	2	US-09-017-628-6

28	507.5	87.0	114	2	US-09-014-880-6	Sequence 6, Appli
29	507.5	87.0	114	4	US-08-450-363-6	Sequence 6, Appli
30	506.5	86.9	260	3	US-08-463-903-2	Sequence 2, Appli
31	506.5	86.9	260	4	US-07-935-695-2	Sequence 2, Appli
32	506	86.8	113	1	US-08-690-102A-6	Sequence 6, Appli
33	506	86.8	113	4	US-09-127-902-6	Sequence 6, Appli
34	506	86.8	113	4	US-09-155-107-6	Sequence 6, Appli
35	506	86.8	113	5	PCT-US95-09641-6	Sequence 6, Appli
36	505.5	86.7	285	3	US-08-463-903-22	Sequence 22, Appli
37	505.5	86.7	285	4	US-07-935-695-22	Sequence 22, Appli
38	504.5	86.5	113	5	PCT-US93-08435-4	Sequence 4, Appli
39	503.5	86.4	113	4	US-09-301-593-6	Sequence 6, Appli
40	503.5	86.4	113	4	US-09-301-593-34	Sequence 34, Appli
41	503.5	86.4	115	4	US-09-025-769B-31	Sequence 31, Appli
42	503.5	86.4	115	4	US-09-025-769B-49	Sequence 49, Appli
43	502.5	86.2	135	3	US-08-812-586-46	Sequence 46, Appli
44	499.5	85.7	113	4	US-09-301-593-2	Sequence 2, Appli
45	499.5	85.7	113	4	US-09-301-593-32	Sequence 32, Appli
46	499.5	85.7	240	4	US-09-301-593-36	Sequence 36, Appli
47	498	85.4	301	2	US-08-661-052-14	Sequence 14, Appli
48	498	85.4	301	4	US-09-188-082-14	Sequence 14, Appli
49	498	85.4	301	4	US-09-364-088-14	Sequence 14, Appli
50	498	85.4	301	4	US-09-102-716-14	Sequence 14, Appli
51	498	85.4	553	2	US-08-661-052-16	Sequence 16, Appli
52	498	85.4	553	4	US-09-188-082-16	Sequence 16, Appli
53	498	85.4	553	4	US-09-364-088-16	Sequence 16, Appli
54	498	85.4	553	4	US-09-102-716-16	Sequence 16, Appli
55	496.5	85.2	114	2	US-08-428-257A-76	Sequence 76, Appli
56	496.5	85.2	114	4	US-07-987-264-16	Sequence 16, Appli
57	492.5	84.5	120	1	US-08-026-320A-4	Sequence 4, Appli
58	490.5	84.1	220	4	US-08-952-235-1	Sequence 1, Appli
59	490.5	84.1	220	4	US-09-669-971-1	Sequence 1, Appli
60	487.5	83.6	113	1	US-08-467-420A-21	Sequence 21, Appli
61	487.5	83.6	113	1	US-08-470-110A-21	Sequence 21, Appli
62	487.5	83.6	113	1	US-08-667-769A-21	Sequence 21, Appli
63	487.5	83.6	113	2	US-08-940-371-21	Sequence 21, Appli
64	487.5	83.6	113	3	US-08-637-647-21	Sequence 21, Appli
65	487.5	83.6	113	5	PCT-US95-17082A-21	Sequence 21, Appli
66	485.5	83.3	113	4	US-09-301-593-4	Sequence 4, Appli
67	485.5	83.3	113	4	US-09-301-593-33	Sequence 33, Appli
68	484	83.0	114	4	US-08-929-856-66	Sequence 66, Appli
69	479.5	82.2	113	5	PCT-US93-11612-5	Sequence 5, Appli
70	479.5	82.2	133	4	US-08-579-378A-2	Sequence 2, Appli
71	479.5	82.2	133	5	PCT-US93-11612-2	Sequence 2, Appli
72	479	82.2	113	4	US-09-155-107-20	Sequence 20, Appli
73	479	82.2	219	1	US-08-353-400-34	Sequence 34, Appli
74	479	82.2	239	1	US-08-353-400-37	Sequence 37, Appli
75	478.5	82.1	133	4	US-09-301-593-24	Sequence 24, Appli
76	478.5	82.1	220	4	US-09-301-593-17	Sequence 17, Appli
77	478.5	82.1	240	4	US-09-301-593-28	Sequence 28, Appli
78	476.5	81.7	113	4	US-08-579-378A-5	Sequence 5, Appli
79	475.5	81.6	113	2	US-08-263-911-2	Sequence 2, Appli
80	475.5	81.6	113	2	US-08-819-033-1	Sequence 1, Appli
81	475.5	81.6	113	4	US-09-025-203-7	Sequence 7, Appli
82	475.5	81.6	113	4	US-09-025-203-13	Sequence 13, Appli
83	475.5	81.6	133	2	US-08-822-028-12	Sequence 12, Appli
84	475.5	81.6	133	3	US-08-479-285-12	Sequence 12, Appli
85	475.5	81.6	275	3	US-08-463-903-8	Sequence 8, Appli
86	475.5	81.6	275	3	US-08-463-903-17	Sequence 17, Appli
87	475.5	81.6	275	4	US-07-935-695-8	Sequence 8, Appli
88	475.5	81.6	275	4	US-07-935-695-17	Sequence 17, Appli
89	475.5	81.6	280	3	US-08-463-903-10	Sequence 10, Appli
90	475.5	81.6	280	4	US-07-935-695-10	Sequence 10, Appli
91	475.5	81.6	282	3	US-08-463-903-12	Sequence 12, Appli
92	475.5	81.6	282	3	US-08-463-903-15	Sequence 15, Appli
93	475.5	81.6	282	4	US-07-935-695-12	Sequence 12, Appli
94	475.5	81.6	282	4	US-07-935-695-15	Sequence 15, Appli
95	475.5	81.6	553	2	US-08-263-911-7	Sequence 7, Appli
96	475.5	81.6	553	2	US-08-263-911-9	Sequence 9, Appli
97	474.5	81.4	242	2	US-08-224-591-14	Sequence 14, Appli
98	474.5	81.4	242	2	US-08-392-338A-23	Sequence 23, Appli
99	474.5	81.4	242	2	US-08-926-789-14	Sequence 14, Appli
100	474.5	81.4	242	3	US-09-166-750-23	Sequence 23, Appli

LENGTH: 112 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-30

Query Match 89.4%; Score 521; DB 1; Length 112;
Best Local Similarity 89.3%; Pred. No. 4.6e-42;
Matches 100; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYQKPGSPKLLIYWASTR 60
Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYQKPGSPKLLIYWASTR 60

Qy 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSTYFGGKLEIK 112
Db 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSTYFGGKLEIK 112

RESULT 7
US-07-916-098A-15
Sequence 15, Application US/07916098A
Patent No. 5871732

GENERAL INFORMATION:
APPLICANT: BURKLY, LINDA C.
APPLICANT: CHISHOLM, PATRICIA L.
APPLICANT: THOMAS, DAVID W.
APPLICANT: ROSA, MARGARET D.
APPLICANT: ROSA, JOSEPH J.
TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
STREET: 10 SOUTH WACKER DRIVE
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/916,098A
FILING DATE: July 24, 1992
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08843
FILING DATE: No. 5871732ember 27, 1991
CLASSIFICATION: 424
APPLICATION NUMBER: 07/618,542
FILING DATE: No. 5871732ember 27, 1990
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: JOHN J. MC DONNELL
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,310-G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1000
TELEFAX: (312) 715-1234
TELEX: 910/221-5317
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-916-098A-15

Query Match 89.4%; Score 521; DB 2; Length 112;
Best Local Similarity 89.3%; Pred. No. 4.6e-42;

Matches 100; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYQKPGSPKLLIYWASTR 60
Db 1 DIVMTQSPDSLAVSGEKVTMCKSSQSVLYSSNKNYLAAYQKPGSPKLLIYWASTR 60

Qy 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSTYFGGKLEIK 112
Db 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSTYFGGKLEIK 112

RESULT 8
PCT-US93-08435-8
Sequence 8, Application PC/TUS9308435
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham, Corporation
APPLICANT: U. S. Government, Secretary of
APPLICANT: the Navy
APPLICANT: the Army
TITLE OF INVENTION: Novel Antibodies for Confering Passive
TITLE OF INVENTION: Immunity Against Infection By a Pathogen in Man
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Box 457, 321 Norristown Road
CITY: Spring House
STATE: PA
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08435
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/941,654
FILING DATE: 09-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: SBC P50107
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 540-9200
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-08435-8

Query Match 89.3%; Score 520.5; DB 5; Length 113;
Best Local Similarity 89.4%; Pred. No. 5.1e-42;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYQKPGSPKLLIYWASTR 60
Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYQKPGSPKLLIYWASTR 60

Qy 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLS-STYFGGKLEIK 112
Db 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLS-STYFGGKLEIK 113

RESULT 9
US-08-828-741B-11
Sequence 11, Application US/08828741B

Patent No. 6043069
GENERAL INFORMATION:
APPLICANT: Koentgen, Frank
APPLICANT: Sues, Gabriele M.
APPLICANT: Tarlington, David M.
APPLICANT: Treutlein, Herbert R.
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,741B
FILING DATE: 26-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10591
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 155 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-828-741B-11

Query Match 89.3%; Score 520.5; DB 3; Length 155;
Best Local Similarity 89.4%; Pred. No. 7.2e-42;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
Qy 1 DIVMTQSPDLSVSLGERVTMCKSSOSVLYSSNKNVLAHYQKPGQSPKLLIYWASTR 60
|||||
30 DIVMTQSPDLSVSLGERVTMCKSSOSVLYSSNKNVLAHYQKPGQSPKLLIYWASTR 89
|||||
61 ESGVDRFSGSGGTDTLTISVQAEDLAVYYCHQYLSS-YTFGGTKLEIK 112
|||||
90 ESGVDRFSGSGGTDTLTISVQAEDLAVYYCHQYVSTPYSFGQTKLEIK 142

RESULT 10
US-09-160-567-11
Sequence 11, Application US/09160567
Patent No. 6326179
GENERAL INFORMATION:
APPLICANT: Koentgen, Frank
APPLICANT: Sues, Gabriele M.
APPLICANT: Tarlington, David M.
APPLICANT: Treutlein, Herbert R.
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/160,567
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/828,741
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10591
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 155 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-160-567-11

Query Match 89.3%; Score 520.5; DB 4; Length 155;
Best Local Similarity 89.4%; Pred. No. 7.2e-42;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
Qy 1 DIVMTQSPDLSVSLGERVTMCKSSOSVLYSSNKNVLAHYQKPGQSPKLLIYWASTR 60
|||||
30 DIVMTQSPDLSVSLGERVTMCKSSOSVLYSSNKNVLAHYQKPGQSPKLLIYWASTR 89
|||||
61 ESGVDRFSGSGGTDTLTISVQAEDLAVYYCHQYLSS-YTFGGTKLEIK 112
|||||
90 ESGVDRFSGSGGTDTLTISVQAEDLAVYYCHQYVSTPYSFGQTKLEIK 142

RESULT 11
US-08-828-741B-6
Sequence 6, Application US/08828741B
Patent No. 6043069
GENERAL INFORMATION:
APPLICANT: Koentgen, Frank
APPLICANT: Sues, Gabriele M.
APPLICANT: Tarlington, David M.
APPLICANT: Treutlein, Herbert R.
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,741B
FILING DATE: 26-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10591

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 342 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-828-741B-6

Query Match      89.3%; Score 520.5; DB 3; Length 342;
Best Local Similarity 89.4%; Pred No. 1.7e-41;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
   |||||
Db 30 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAHYQKPGQPPKLLIYWASTR 89
   |||||
QY 61 ESGVDPDRFSGSGGTFDTLTSSVQAEADLAHYCHQYLSS-YTFGGGKLEIK 112
   |||||
Db 90 ESGVDPDRFSGSGGTFDTLTSSVQAEADVAHYCOQYSTPYFSGGKLEIK 142
   |||||

RESULT 13
US-08-828-741B-4
; Sequence 4, Application US/08828741B
; Patent No. 6043069
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828.741B
; FILING DATE: 26-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-828-741B-4

Query Match      89.3%; Score 520.5; DB 3; Length 495;
Best Local Similarity 89.4%; Pred No. 2.5e-41;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
   |||||
Db 30 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAHYQKPGQPPKLLIYWASTR 89
   |||||
QY 61 ESGVDPDRFSGSGGTFDTLTSSVQAEADLAHYCHQYLSS-YTFGGGKLEIK 112
   |||||
Db 90 ESGVDPDRFSGSGGTFDTLTSSVQAEADVAHYCOQYSTPYFSGGKLEIK 142
   |||||

RESULT 14
US-09-160-567-4
; Sequence 4, Application US/09160567
; Patent No. 6326179

Query Match      89.3%; Score 520.5; DB 4; Length 342;
Best Local Similarity 89.4%; Pred No. 1.7e-41;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
   |||||
Db 30 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAHYQKPGQPPKLLIYWASTR 89
   |||||
QY 61 ESGVDPDRFSGSGGTFDTLTSSVQAEADLAHYCHQYLSS-YTFGGGKLEIK 112
   |||||
Db 90 ESGVDPDRFSGSGGTFDTLTSSVQAEADVAHYCOQYSTPYFSGGKLEIK 142
   |||||

US-09-160-567-6
; Sequence 6, Application US/09160567
; Patent No. 6326179
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/160,567
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828.741
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 342 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-160-567-6
```

```

: GENERAL INFORMATION:
: APPLICANT: Koentgen, Frank
: APPLICANT: Sues, Gabriele M.
: APPLICANT: Tarlinton, David M.
: APPLICANT: Treutlein, Herbert R.
: TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
: TITLE OF INVENTION: PRODUCING SAME
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
: STREET: 400 Garden City Plaza
: CITY: Garden City
: STATE: New York
: COUNTRY: United States of America
: ZIP: 11530
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/160.567
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/828.741
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Digiglio, Frank S.
: REGISTRATION NUMBER: 31,346
: REFERENCE/DOCKET NUMBER: 10591
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (516) 742-4343
: TELEFAX: (516) 742-4366
: TELEX: 230 901 SANS UR
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 495 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-160-567-4

```

```

Query Match 89.3%; Score 520.5; DB 4; Length 495;
Best Local Similarity 89.4%; Pred. No. 2.5e-41;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

```

```

1 DIVMTQSPDSLAVSLGERVTNCKSSQSVLYSSNOKNYLAWYQOKPGSPKLLIYWASTR 60
|||||
30 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNOKNYLAWYQOKPGPPKLLIYWASTR 89
|||||
61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLSS-YTFGGTKLEIK 112
|||||
90 ESGVPDRFSGSGGTDFLTITSSVQAEDVAVVYCOQYVSTYPSFGQTKLEIK 142
|||||

```

```

RESULT 15
US-07-916-098A-56
: Sequence 56, Application US/07916098A
: Patent No. 5871732
: GENERAL INFORMATION:
: APPLICANT: BURKLY, LINDA C.
: APPLICANT: CHISHOLM, PATRICIA L.
: APPLICANT: THOMAS, DAVID W.
: APPLICANT: ROSA, MARGARET D.
: APPLICANT: ROSA, JOSEPH J.
: TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
: TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
: NUMBER OF SEQUENCES: 61
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
: STREET: 10 SOUTH WACKER DRIVE
: CITY: CHICAGO

```

```

: STATE: ILLINOIS
: COUNTRY: U.S.A.
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORD PERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/916.098A
: FILING DATE: July 24, 1992
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US91/08843
: FILING DATE: No. 5871732ember 27, 1991
: CLASSIFICATION: 424
: APPLICATION NUMBER: 07/618.542
: FILING DATE: No. 5871732ember 27, 1990
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: JOHN J. MC DONNELL
: REGISTRATION NUMBER: 26,949
: REFERENCE/DOCKET NUMBER: 92,310-G
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 715-1000
: TELEFAX: (312) 715-1234
: TELEX: 910/221-5317
: INFORMATION FOR SEQ ID NO: 56:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 241 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-916-098A-56

```

```

Query Match 88.3%; Score 515; DB 2; Length 241;
Best Local Similarity 88.4%; Pred. No. 3.8e-41;
Matches 99; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 1 DIVMTQSPDSLAVSLGERVTNCKSSQSVLYSSNOKNYLAWYQOKPGSPKLLIYWASTR 60
|||||
DB 23 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNOKNYLAWYQOKPGPPKLLIYWASTR 82
|||||
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLSSYTFGGTKLEIK 112
|||||
DB 83 ESGVPDRFSGSGGTDFLTITSSVQAEDVAVVYCOQYVSYRTFGRGTLEIK 134
|||||

```

```

RESULT 16
US-08-525-539A-80
: Sequence 80, Application US/08525539A
: Patent No. 6309636
: GENERAL INFORMATION:
: APPLICANT: DO COUTO, FERNANDO J.R.
: APPLICANT: CERIANI, ROBERTO L.
: APPLICANT: PETERSON, JERRY A.
: TITLE OF INVENTION: RECOMBINANT PEPTIDES DERIVED FROM THE
: TITLE OF INVENTION: MC3 ANTI-BA46 ANTIBODY, METHODS OF USE THEREOF, AND
: TITLE OF INVENTION: METHODS OF HUMANIZING ANTIBODY PEPTIDES
: NUMBER OF SEQUENCES: 81
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORRISON & FOERSTER
: STREET: 755 Page Mill Road
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304-1018
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:

```

61 ESGVPDRFSGSGGTDFTLTIS~~SS~~VQAEDLAVYCHQYLS-SYTFGGGTKLEIK 112

Db 83 ESGVDRFSGSGCTDFTLTISLSQAEDVAVYYCQYYSYPLTFGGGKVVIK 135

RESULT 20

US-07-935-695-6
 ; Sequence 6, Application US/07935695
 ; Patent No. 6329507
 ; GENERAL INFORMATION:
 ; APPLICANT: Mezes, Peter S.
 ; APPLICANT: Richard, Ruth A.
 ; APPLICANT: Affholter, Joseph A.
 ; APPLICANT: Kotite, Nicolas J.
 ; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
 ; FILE REFERENCE: 40224A US
 ; CURRENT APPLICATION NUMBER: US/07/935.695
 ; PRIORITY FILING DATE: 1992-08-21
 ; PRIOR APPLICATION NUMBER: US 08/463,903
 ; PRIOR FILING DATE: 1995-06-05
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: MS-Word for Windows, Ver. 7.0
 ; ID NO 6
 ; LENGTH: 275
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; NAME/KEY: Hum4VL-VH scFv from pSCFV UHH
 ; LOCATION: 1..275
 ; OTHER INFORMATION: :
 US-07-935-695-6

Query Match 87.9%; Score 512.5; DB 4; Length 275;
 Best Local Similarity 88.5%; Pred. No. 7.6e-41;
 Matches 100; Conservative 4; Mismatches 8; Indels 1; Gaps 1;
 Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSOSVLYSSNKNYLAAYQKPGOSPKLLIYWASTR 60
 Db 23 DIVMTQSPDSLAVSLGERATINCKSSOSVLYSSNKNYLAAYQKPGOPPKLLIYWASTR 82
 Qy 61 ESGVDRFSGSGCTDFTLTISLSQAEDVAVYYCQYYSYPLTFGGGKLEIK 112
 Db 83 ESGVDRFSGSGCTDFTLTISLSQAEDVAVYYCQYYSYPLTFGGGKVVIK 135

RESULT 21

US-09-025-769B-17
 ; Sequence 17, Application US/09025769B
 ; Patent No. 6300064
 ; GENERAL INFORMATION:
 ; APPLICANT: Knappik, Achim
 ; APPLICANT: Pack, Peter
 ; APPLICANT: Ilag, Vic
 ; APPLICANT: Ge, Liming
 ; APPLICANT: Moroney, Simon
 ; APPLICANT: Plueckthun, Andreas
 ; TITLE OF INVENTION: Protein/(Poly)peptide libraries
 ; NUMBER OF SEQUENCES: 373
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
 ; STREET: 1251 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10021
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/025,769B
 ; FILING DATE: 18-FEB-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: EP 95 11 3021.0

FILING DATE: 18-AUG-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: James F. Haley, Jr., Esq.
 REGISTRATION NUMBER: 27,794
 REFERENCE/DOCKET NUMBER: MORPHO/5
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)596-9000
 TELEFAX: (212)596-9090
 INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 114 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-025-769B-17
 Query Match 87.7%; Score 511.5; DB 4; Length 114;
 Best Local Similarity 88.5%; Pred. No. 3.6e-41;
 Matches 100; Conservative 5; Mismatches 7; Indels 1; Gaps 1;
 Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSOSVLYSSNKNYLAAYQKPGOSPKLLIYWASTR 60
 Db 1 DIVMTQSPDSLAVSLGERATINCKSSOSVLYSSNKNYLAAYQKPGOPPKLLIYWASTR 60
 Qy 61 ESGVDRFSGSGCTDFTLTISLSQAEDVAVYYCQYYSYPLTFGGGKLEIK 112
 Db 61 ESGVDRFSGSGCTDFTLTISLSQAEDVAVYYCQYYSYPLTFGGGKVEIK 113

RESULT 22

US-08-483-749A-16
 ; Sequence 16, Application US/08483749A
 ; Patent No. 6054561
 ; GENERAL INFORMATION:
 ; APPLICANT: RING, DAVID B.
 ; TITLE OF INVENTION: ANTIGEN-BINDING SITES OF ANTIBODY
 ; TITLE OF INVENTION: MOLECULES SPECIFIC FOR CANCER ANTIGENS
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CHIRON CORPORATION
 ; STREET: INTELLECTUAL PROPERTY - R440, PO BOX 8097
 ; CITY: EMERYVILLE
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94662-8097
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/483,749A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: SAVERIDE, PAUL B.
 ; REGISTRATION NUMBER: 36,914
 ; REFERENCE/DOCKET NUMBER: 0508.008
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (510) 601-2585
 ; TELEFAX: (510) 655-3542
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 113 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-483-749A-16

Query Match 87.6%; Score 510.5; DB 3; Length 113;
 Best Local Similarity 86.7%; Pred. No. 4.4e-41;
 Matches 98; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 60
Db 1 ELVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 60
QY 61 ESGVDPFRFSGSGDFTLTISVQAEDLAVYCHQYLS-SYTFGGGKLEIK 112
Db 61 ESGVDPFRFSGSGDFTLTISVQAEDLAVYCHQYLS-SYTFGGGKLEIK 113

RESULT 23
US-08-463-903-20
Sequence 20, Application US/08463903
Patent No. 6071515
GENERAL INFORMATION:
APPLICANT: Mezes, Peter S.
APPLICANT: Richard, Ruth A.
APPLICANT: Affholter, Joseph A.
APPLICANT: Kotite, Nicolas J.
TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
FILE REFERENCE: 40224A US
CURRENT APPLICATION NUMBER: US/08/463,903
CURRENT FILING DATE: 1995-06-05
EARLIER APPLICATION NUMBER: US 07/935,695
EARLIER FILING DATE: 1992-08-21
NUMBER OF SEQ ID NOS: 102
SOFTWARE: MS-Word for Windows, Ver. 7.0
SEQ ID NO 20
LENGTH: 171
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: Hum4 VL-UNIHOPE linker-FLAG peptide construct in PATDFLAG
LOCATION: 1..171
US-08-463-903-20

Query Match 87.6%; Score 510.5; DB 3; Length 171;
Best Local Similarity 88.5%; Pred. No. 6.9e-41;
Matches 100; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 60
Db 23 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 82
QY 61 ESGVDPFRFSGSGDFTLTISVQAEDLAVYCHQYLS-SYTFGGGKLEIK 112
Db 83 ESGVDPFRFSGSGDFTLTISVQAEDLAVYCHQYLS-SYTFGGGKLEIK 135

JUL 24
US-07-935-695-20
Sequence 20, Application US/07935695
Patent No. 6329507
GENERAL INFORMATION:
APPLICANT: Mezes, Peter S.
APPLICANT: Richard, Ruth A.
APPLICANT: Affholter, Joseph A.
APPLICANT: Kotite, Nicolas J.
TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
FILE REFERENCE: 40224A US
CURRENT APPLICATION NUMBER: US/07/935,695
CURRENT FILING DATE: 1992-08-21
PRIOR APPLICATION NUMBER: US 08/463,903
PRIOR FILING DATE: 1995-06-05
NUMBER OF SEQ ID NOS: 102
SOFTWARE: MS-Word for Windows, Ver. 7.0
SEQ ID NO 20
LENGTH: 171
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: Hum4 VL-UNIHOPE linker-FLAG peptide construct in PATDFLAG
LOCATION: 1..171

OTHER INFORMATION:
US-07-935-695-20

Query Match 87.6%; Score 510.5; DB 4; Length 171;
Best Local Similarity 88.5%; Pred. No. 6.9e-41;
Matches 100; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 60
Db 23 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNKNYLAHYQKPGSPKLLIYWASTR 82
QY 61 ESGVDPFRFSGSGDFTLTISVQAEDLAVYCHQYLS-SYTFGGGKLEIK 112
Db 83 ESGVDPFRFSGSGDFTLTISVQAEDLAVYCHQYLS-SYTFGGGKLEIK 135

RESULT 25

US-08-360-125-6
Sequence 6, Application US/08360125
Patent No. 5767246
GENERAL INFORMATION:
APPLICANT: Saiko HOSOKAWA
APPLICANT: Toshiaki TAGAWA
APPLICANT: Yoko HIRAKAWA
APPLICANT: No. 5767246.hiko ITO
APPLICANT: Kazuhiro NAGAIKE
TITLE OF INVENTION: Human Monoclonal Antibody
SPECIFICALLY BINDING TO SURFACE ANTIGEN OF CANCER
TITLE OF INVENTION: Cell Membrane
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,125
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/905,534
FILING DATE: June 29, 1992
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:


```

DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE: Hybridoma producing human
CELL TYPE: antibody GAH
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-360-125-6

```

```

Query Match      87.0%; Score 507.5; DB 1; Length 114;
Best Local Similarity 86.7%; Pred No. 8.5e-41;
Matches 98; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTNCKSSQSVLYSSNKNYLAHYQOKPGQSPKLLIYWASTR 60
   |||||
Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSNKNKYLAHYQOKPGQSPKLLIYWASTR 60
   |||||

Qy 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHOYLSS-YTFGGGTKLEIK 112
   |||||
Db 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHOYLSS-YTFGGGTKLEIK 113
   |||||

```

Search completed: June 23, 2003, 14:04:52
time : 7.55319 secs

93 429.5 73.7 133 9 US-09-874-141-47 Sequence 47, Appl
94 427.5 73.3 107 9 US-09-874-141-4 Sequence 4, Appl
95 425.5 73.0 151 9 US-09-187-693-58 Sequence 58, Appl
96 425.5 73.0 159 9 US-09-187-693-66 Sequence 66, Appl
97 424.5 72.8 131 9 US-09-873-461-58 Sequence 58, Appl
98 419.5 72.0 107 9 US-09-956-206A-81 Sequence 81, Appl
99 419.5 72.0 111 9 US-10-078-958-16 Sequence 16, Appl
100 419.5 72.0 127 9 US-09-956-206A-65 Sequence 65, Appl

ALIGNMENTS

RESULT 1

US-10-056-052-18

; Sequence 18, Application US/10056052
; Publication No. US20030099656A1

; GENERAL INFORMATION:

; APPLICANT: HUTCHINS, Jeff T

; APPLICANT: DOMANSKI, Paul

; APPLICANT: PATEL, Pratiksha

; APPLICANT: HALL, Andrea

; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN

; FILE REFERENCE: P07069US04/BAS

; CURRENT FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/308,116

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/298,413

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: 60/274,611

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: 60/264,072

; PRIOR FILING DATE: 2001-01-26

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 18

; LENGTH: 112

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-10-056-052-18

Query Match 100.0%; Score 583; DB 9; Length 112;
Best Local Similarity 100.0%; Pred. No. 3.1e-37;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60

1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60

61 ESGVDPDRFGSGSGTDFTLTISVQAEADLAVYCHQYLSSTYFGGTTKLEIK 112

61 ESGVDPDRFGSGSGTDFTLTISVQAEADLAVYCHQYLSSTYFGGTTKLEIK 112

RESULT 2

US-10-056-052-10

; Sequence 10, Application US/10056052

; Publication No. US20030099656A1

; GENERAL INFORMATION:

; APPLICANT: HUTCHINS, Jeff T

; APPLICANT: DOMANSKI, Paul

; APPLICANT: PATEL, Pratiksha

; APPLICANT: HALL, Andrea

; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN

; FILE REFERENCE: P07069US04/BAS

; CURRENT FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/308,116

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/298,413

; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-10

Query Match 94.9%; Score 553; DB 9; Length 112;

Best Local Similarity 93.8%; Pred. No. 5.5e-35;

Matches 105; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60

Db 1 NIMTQSPSSSLAVSAGEKVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60

QY 61 ESGVDPDRFGSGSGTDFTLTISVQAEADLAVYCHQYLSSTYFGGTTKLEIK 112

Db 61 ESGVDPDRFGSGSGTDFTLTISVQAEADLAVYCHQYLSSTYFGGTTKLEIK 112

RESULT 3

US-10-056-052-6

; Sequence 6, Application US/10056052

; Publication No. US20030099656A1

; GENERAL INFORMATION:

; APPLICANT: HUTCHINS, Jeff T

; APPLICANT: DOMANSKI, Paul

; APPLICANT: PATEL, Pratiksha

; APPLICANT: HALL, Andrea

; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN

; FILE REFERENCE: P07069US04/BAS

; CURRENT FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/308,116

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/298,413

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: 60/274,611

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: 60/264,072

; PRIOR FILING DATE: 2001-01-26

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 112

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-10-056-052-6

Query Match 93.5%; Score 545; DB 9; Length 112;

Best Local Similarity 92.0%; Pred. No. 2.2e-34;

Matches 103; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60

Db 1 NIMTQSPSSSLAVSAGEKVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60

QY 61 ESGVDPDRFGSGSGTDFTLTISVQAEADLAVYCHQYLSSTYFGGTTKLEIK 112

Db 61 ESGVDPDRFGSGSGTDFTLTISVQAEADLAVYCHQYLSSTYFGGTTKLEIK 112

RESULT 4

US-10-056-052-14

; Sequence 14, Application US/10056052

; Publication No. US20030099656A1

```

: GENERAL INFORMATION:
: APPLICANT: PATTI, Joseph M
: APPLICANT: HUTCHINS, Jeff T
: APPLICANT: DOMANSKI, Paul
: APPLICANT: PATEL, Pratiksha
: APPLICANT: HALL, Andrea
: TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
: FILE REFERENCE: P070690504/BAS
: CURRENT APPLICATION NUMBER: US/10/056,052
: CURRENT FILING DATE: 2002-04-19
: PRIOR APPLICATION NUMBER: 60/308,116
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/298,413
: PRIOR FILING DATE: 2001-06-18
: PRIOR APPLICATION NUMBER: 60/274,611
: PRIOR FILING DATE: 2001-03-12
: PRIOR APPLICATION NUMBER: 60/264,072
: PRIOR FILING DATE: 2001-01-26
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 14
: LENGTH: 112
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
: US-10-056-052-14

Query Match          93.3%; Score 544; DB 9; Length 112;
Best Local Similarity 92.0%; Pred. No. 2.6e-34;
Matches 103; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 DIVWTSPLSLVSLGERTVMTNCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60
Db 1 NIMWTSPLSLVSLGERTVMTNCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60
QY 61 ESCVPRFSGSGTDTFTLTISVQAEEDLAVYYCHQYLSSTYFGGKLEIK 112
Db 61 ESCVPRFSGSGTDTFTLTISVQAEEDLAVYYCHQYLSSTYFGGKLEIK 112

RESULT 5
US-09-894-839-2
: Sequence 2, Application US/09894839
: Publication No. US20030035800A1
: GENERAL INFORMATION:
: APPLICANT: LEUNG, Hans
: APPLICANT: OU, Zhengxing
: TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
: FILE REFERENCE: 018733/1049
: CURRENT APPLICATION NUMBER: US/09/894,839
: CURRENT FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 09/155,107
: PRIOR FILING DATE: 1998-11-17
: PRIOR APPLICATION NUMBER: US 60/013,709
: PRIOR FILING DATE: 1996-03-20
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 2
: LENGTH: 113
: TYPE: PRT
: ORGANISM: Murinae gen. sp.
: US-09-894-839-2

Query Match          90.4%; Score 527; DB 9; Length 113;
Best Local Similarity 88.4%; Pred. No. 5e-33;
Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 DIVWTSPLSLVSLGERTVMTNCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60
Db 1 DIQWTSPLSLVSLGERTVMTNCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60
QY 61 ESCVPRFSGSGTDTFTLTISVQAEEDLAVYYCHQYLSSTYFGGKLEIK 112
Db 61 ESCVPRFSGSGTDTFTLTISVQAEEDLAVYYCHQYLSSTYFGGKLEIK 112

RESULT 6
US-09-988-013A-2
: Sequence 2, Application US/09988013A
: Publication No. US20030103979A1
: GENERAL INFORMATION:
: APPLICANT: LEUNG, Hans
: APPLICANT: HANSEN, Hans
: TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
: TITLE OF INVENTION: CELL LYMPHOMA AND LEUKEMIA CELLS
: FILE REFERENCE: 018733/1082
: CURRENT APPLICATION NUMBER: US/09/988,013A
: CURRENT FILING DATE: 2002-10-29
: PRIOR APPLICATION NUMBER: US 09/741,843
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/127,902
: PRIOR FILING DATE: 1998-08-03
: PRIOR APPLICATION NUMBER: US 08/690,102
: PRIOR FILING DATE: 1996-07-06
: PRIOR APPLICATION NUMBER: US 08/289,576
: PRIOR FILING DATE: 1994-08-12
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 2
: LENGTH: 113
: TYPE: PRT
: ORGANISM: Murinae gen. sp.
: US-09-988-013A-2

Query Match          90.4%; Score 527; DB 9; Length 113;
Best Local Similarity 88.4%; Pred. No. 5e-33;
Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 DIVWTSPLSLVSLGERTVMTNCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60
Db 1 DIQWTSPLSLVSLGERTVMTNCKSSQSVLYSSNQKNYLAWSYQKQSPKLLIYWASTR 60
QY 61 ESCVPRFSGSGTDTFTLTISVQAEEDLAVYYCHQYLSSTYFGGKLEIK 112
Db 61 ESCVPRFSGSGTDTFTLTISVQAEEDLAVYYCHQYLSSTYFGGKLEIK 112

RESULT 7
US-09-741-843-2
: Sequence 2, Application US/09741843
: Patent No. US20020102254A1
: GENERAL INFORMATION:
: APPLICANT: LEUNG, Hans
: APPLICANT: HANSEN, Hans
: TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
: TITLE OF INVENTION: AND LEUKEMIA CELLS
: FILE REFERENCE: 018733/0996
: CURRENT APPLICATION NUMBER: US/09/741,843
: CURRENT FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/127,902
: PRIOR FILING DATE: 1998-08-03
: PRIOR APPLICATION NUMBER: US 08/690,102
: PRIOR FILING DATE: 1996-07-06
: PRIOR APPLICATION NUMBER: US 08/289,576
: PRIOR FILING DATE: 1994-08-12
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 2
: LENGTH: 113
: TYPE: PRT
: ORGANISM: Murinae gen. sp.
: US-09-741-843-2

Query Match          90.4%; Score 527; DB 10; Length 113;
Best Local Similarity 88.4%; Pred. No. 5e-33;
Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
```



```

: Sequence 6, Application US/0995529
: Publication No. US20030099655A1
: GENERAL INFORMATION:
: APPLICANT: Watkins, Jeffrey D.
: APPLICANT: Huse, William D.
: APPLICANT: Tang, Ying
: TITLE OF INVENTION: Humanized Collagen Antibodies and
: TITLE OF INVENTION: Related Methods
: FILE REFERENCE: P-IX 4976
: CURRENT APPLICATION NUMBER: US/09/995,529
: CURRENT FILING DATE: 2001-11-26
: NUMBER OF SEQ ID NOS: 358
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 113
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-995-529-6

Query Match 87.6%; Score 510.5; DB 9; Length 113;
Best Local Similarity 88.5%; Pred. No. 8.7e-32;
Matches 100; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNKNYLAHYQKPGQPPKLLIYWASTR 60
Qy 61 ESGVPRFSGSGGDTFTLTISVQAEADLVAVYCHQ-YLSSVTFGGGTGLEIK 112
Db 61 ESGVPRFSGSGGDTFTLTISLQAEADVAVYCCQDHSYPTFGGTGLEIK 113

RESULT 16
US-09-810-502-38
: Sequence 38, Application US/09810502
: Patent No. US20020034765A1
: GENERAL INFORMATION:
: APPLICANT: Padlan, Eduardo A.
: Daugherty, Bruce L.
: Mark, George E.
: TITLE OF INVENTION: A METHOD FOR REDUCING THE IMMUNOGENICITY
: OF ANTIBODY VARIABLE DOMAINS
: NUMBER OF SEQUENCES: 38
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: P.O. Box 2000, 126 E. Lincoln Ave.
: CITY: Rahway
: STATE: NJ
: COUNTRY: USA
: ZIP: 07065-0907
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows
: SOFTWARE: FastSeq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/810,502
: FILING DATE: 16-Mar-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/905,280
: FILING DATE: 01-Aug-1997
: APPLICATION NUMBER: 08/609,218
: FILING DATE: 01-Mar-1996
: APPLICATION NUMBER: 08/109,187
: FILING DATE: 19-Aug-1993
: APPLICATION NUMBER: 07/702,217
: FILING DATE: 17-May-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Hand, J. Mark
: REGISTRATION NUMBER: 36,545
: REFERENCE/DOCKET NUMBER: 18410CC
: TELECOMMUNICATION INFORMATION:

: Sequence 6, Application US/0995529
: Publication No. US20030099655A1
: TELEPHONE: 732-594-3905
: TELEFAX: 732-594-4720
: TELEX: <Unknown>
: INFORMATION FOR SEQ ID NO: 38:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 114 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-810-502-38

Query Match 87.0%; Score 507.5; DB 10; Length 114;
Best Local Similarity 87.6%; Pred. No. 1.5e-31;
Matches 99; Conservative 6; Mismatches 7; Indels 1; Gaps 1

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSSSLAVSLGERATINCKSSQSVLYSSNKNYLAHYQKPGQPPKLLIYWASTR 60
Qy 61 ESGVPRFSGSGGDTFTLTISVQAEADLVAVYCHQYLS-YTFGGGTGLEIK 112
Db 61 ESGVPRFSGSGGDTFTLTISLQAEADVAVYCCQYIYSPYSGGTGLEIK 113

RESULT 17
US-10-121-464-6
: Sequence 6, Application US/10121464
: Publication No. US20030103968A1
: GENERAL INFORMATION:
: APPLICANT: Boehringer Ingelheim International GmbH
: APPLICANT: Boehringer Ingelheim Pharmaceuticals, Inc.
: TITLE OF INVENTION: Cancer treatment by using FAP-alpha specific antibodies
: FILE REFERENCE: 1-1203ff
: CURRENT APPLICATION NUMBER: US/10/121,464
: CURRENT FILING DATE: 2002-11-15
: PRIOR APPLICATION NUMBER: US 60/283,868
: PRIOR FILING DATE: 2001-04-12
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 113
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Antibody
: OTHER INFORMATION: sequence
US-10-121-464-6

Query Match 86.4%; Score 503.5; DB 9; Length 113;
Best Local Similarity 86.7%; Pred. No. 2.9e-31;
Matches 98; Conservative 5; Mismatches 9; Indels 1; Gaps 1

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNKNYLAHYQKPGQPPKLLIYWASTR 60
Qy 61 ESGVPRFSGSGGDTFTLTISVQAEADLVAVYCHQYLS-SYTFGGGTGLEIK 112
Db 61 ESGVPRFSGSGGDTFTLTISLQAEADVAVYCCQYFYSPYTFGGGTGLEIK 113

RESULT 18
US-10-125-687-11
: Sequence 11, Application US/10125687
: Publication No. US20030054407A1
: GENERAL INFORMATION:
: APPLICANT: Luo, Peter
: TITLE OF INVENTION: STRUCTURE-BASED CONSTRUCTION OF HUMAN ANTIBODY LIBRARY
: FILE REFERENCE: 26050-705
: CURRENT APPLICATION NUMBER: US/10/125,687
: CURRENT FILING DATE: 2002-04-17

```



```
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human consensus antibody light chain variable region
US-10-125-687-11

Query Match      86.4%; Score 503.5; DB 9; Length 114;
Best Local Similarity 86.7%; Pred. No. 2.9e-31;
Matches 98; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNKYLAWYQKPGSPKLLIYWASTR 60
Db 1 DIVMTQSPDLSAVSLGERATINCRSSQSVLYSSNKNKYLAWYQKPGPPKLLIYWASTR 60
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQ-YLSSYTFGGGTKLEIK 112
61 ESGVPDRFSGSGGTDFLTITSSVQAEADVAVYYCQOHYTPPTFGQGTKVEIK 113

RESULT 19
US-10-025-687-11
; Sequence 11, Application US/10025687
; Patent No. US20020142255A1
; GENERAL INFORMATION:
; APPLICANT: Luo, Peter
; TITLE OF INVENTION: STRUCTURE-BASED CONSTRUCTION OF HUMAN ANTIBODY LIBRARY
; FILE REFERENCE: 26050-705
; CURRENT APPLICATION NUMBER: US/10/025.687
; CURRENT FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human consensus antibody light chain variable region
US-10-025-687-11

Query Match      86.4%; Score 503.5; DB 12; Length 114;
Best Local Similarity 86.7%; Pred. No. 2.9e-31;
Matches 98; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNKYLAWYQKPGSPKLLIYWASTR 60
1 DIVMTQSPDLSAVSLGERATINCRSSQSVLYSSNKNKYLAWYQKPGPPKLLIYWASTR 60
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQ-YLSSYTFGGGTKLEIK 112
61 ESGVPDRFSGSGGTDFLTITSSVQAEADVAVYYCQOHYTPPTFGQGTKVEIK 113

RESULT 20
US-10-171-452A-1
; Sequence 1, Application US/10171452A
; Publication No. US20030108518A1
; GENERAL INFORMATION:
; APPLICANT: Frewin, Mark
; APPLICANT: Waldmann, Herman
; APPLICANT: Gorman, Scott
; APPLICANT: Hale, Geoff
; APPLICANT: Rao, Patricia
; APPLICANT: Kornaga, Tadeusz
; APPLICANT: Ringler, Douglas
; APPLICANT: Cobbold, Stephen
; APPLICANT: Winsor-Hines, Dawn
; TITLE OF INVENTION: TRX1 Antibody and Uses Therefor
; FILE REFERENCE: 695458-59
; CURRENT APPLICATION NUMBER: US/10/171.452A
```

```
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: US60/373,471
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US60/373,470
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US60/345,194
; PRIOR FILING DATE: 2002-10-19
; PRIOR APPLICATION NUMBER: GB0122724.8
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: GB0114517.6
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 60
; SEQ ID NO 1
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-452A-1

Query Match      86.2%; Score 502.5; DB 9; Length 135;
Best Local Similarity 86.7%; Pred. No. 4e-31;
Matches 98; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNKYLAWYQKPGSPKLLIYWASTR 60
Db 21 DIVMTQSPDLSAVSLGERATINCRSSQSVLYSSNKNKYLAWYQKPGPPKLLIYWASTR 80
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQYLSS-YTFGGGTKLEIK 112
61 ESGVPDRFSGSGGTDFLTITSSVQAEADVAVYYCQYIYTPPMFGQGTKVEIK 133

RESULT 21
US-09-894-839-6
; Sequence 6, Application US/09894839
; Publication No. US2003003800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; APPLICANT: OU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894.839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-6

Query Match      85.8%; Score 500; DB 9; Length 113;
Best Local Similarity 81.2%; Pred. No. 5.3e-31;
Matches 91; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNKYLAWYQKPGSPKLLIYWASTR 60
Db 1 DIQLTQSPSSLASVSGDRVTMCKSSQSVLYSANHKNYLAWYQKPGKAPKLLIYWASTR 60
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQYLSSYTFGGGTKLEIK 112
61 ESGVPDRFSGSGGTDFLTITSSVQAEADVAVYYCHQYLSSYTFGGGTKVQIK 112

RESULT 22
US-09-988-013A-6
; Sequence 6, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
```

APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans
TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYMPHOMA AND LEUKEMIA CELLS
FILE REFERENCE: 18733/1082
CURRENT APPLICATION NUMBER: US/09/988,013A
CURRENT FILING DATE: 2002-10-29
PRIOR APPLICATION NUMBER: US 09/741,843
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/127,902
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 08/690,102
PRIOR FILING DATE: 1996-07-06
PRIOR APPLICATION NUMBER: US 08/289,576
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 113
TYPE: PRT
ORGANISM: Homo sapiens
US-988-013A-6

Query Match 85.8%; Score 500; DB 9; Length 113;
Best Local Similarity 81.2%; Pred. No. 5.3e-31;
Matches 91; Conservative 13; Mismatches 8; Indels 0; Gaps 0;
Qy 1 DIVMTQSPDLSAVSLGERVTWCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIQLTQSPSSLSASVGDRTWCKSSQSVLYSSNOKNYLAWYQKPGKAPKLLIYWASTR 60
Qy 61 ESGVDPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLSSTYFGGKLEIK 112
Db 61 ESGVPSRFSGSGGTDFLTITSSLPQEDIAITYCHQYLSSTWTFGGTKVQIK 112

RESULT 23
US-09-741-843-6
Sequence 6, Application US/09741843
Patent No. US20020102254A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans
TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYMPHOMA AND LEUKEMIA CELLS
FILE REFERENCE: 018733/0996
CURRENT APPLICATION NUMBER: US/09/741,843
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/127,902
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 08/690,102
PRIOR FILING DATE: 1996-07-06
PRIOR APPLICATION NUMBER: US 08/289,576
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 113
TYPE: PRT
ORGANISM: Homo sapiens
US-09-741-843-6

Query Match 85.8%; Score 500; DB 10; Length 113;
Best Local Similarity 81.2%; Pred. No. 5.3e-31;
Matches 91; Conservative 13; Mismatches 8; Indels 0; Gaps 0;
Qy 1 DIVMTQSPDLSAVSLGERVTWCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIQLTQSPSSLSASVGDRTWCKSSQSVLYSSNOKNYLAWYQKPGKAPKLLIYWASTR 60
Qy 61 ESGVDPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLSSTYFGGKLEIK 112
Db 61 ESGVPSRFSGSGGTDFLTITSSLPQEDIAITYCHQYLSSTWTFGGTKVQIK 112

RESULT 24

US-10-121-464-2
Sequence 2, Application US/10121464
Publication No. US20030103968A1
GENERAL INFORMATION:
APPLICANT: Boehringer Ingelheim International GmbH
APPLICANT: Boehringer Ingelheim Pharmaceuticals, Inc.
TITLE OF INVENTION: Cancer treatment by using FAP-alpha specific antibodies
FILE REFERENCE: 1-1203ff
CURRENT APPLICATION NUMBER: US/10/121,464
CURRENT FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: US 60/283,868
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 2
LENGTH: 113
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Antibody
OTHER INFORMATION: sequence
US-10-121-464-2

Query Match 85.7%; Score 499.5; DB 9; Length 113;
Best Local Similarity 85.8%; Pred. No. 5.8e-31;
Matches 97; Conservative 6; Mismatches 9; Indels 1; Gaps 1;
Qy 1 DIVMTQSPDLSAVSLGERVTWCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Qy 61 ESGVDPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLS-SYTFGGGKLEIK 112
Db 61 ESGVDPDRFSGSGGTDFLTITSSLOAEDVAVVYCOQYFSYPLTFGGGKVEIK 113

RESULT 25

US-09-828-708-1
Sequence 1, Application US/09828708
Patent No. US20020146753A1
GENERAL INFORMATION:
APPLICANT: Ditzel, H.
APPLICANT: Burton, D.
APPLICANT: Schaller, M.
TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their part
FILE REFERENCE: 1361.005US1
CURRENT APPLICATION NUMBER: US/09/828,708
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 123
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 109
TYPE: PRT
ORGANISM: Homo sapiens
US-09-828-708-1

Query Match 84.7%; Score 494; DB 10; Length 109;
Best Local Similarity 88.6%; Pred. No. 1.5e-30;
Matches 93; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
Qy 8 PDSLAVSLGERVTWCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTRESGVDR 67
Db 1 PDSLAVSLGERATINCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTRESGVDR 60
Qy 68 FSGSGGTDFLTITSSVQAEDLAVVYCHQYLSSTYFGGKLEIK 112
Db 61 FSGSGGTDFLTITSSLOAEDVAVVYCOQYDSYTFGGGKLEIK 105

Search completed: June 23, 2003, 14:19:13
Job time : 16.8865 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:31 ; Search time 7.5461 Seconds
(without alignments)
1426.837 Million cell updates

Title: US-10-056-052A-18
 Perfect score: 583
 Sequence: 1 DIVMTQSPDLSAVSLGERVT.....YCHOYLSSYTFGGGTKLEIK 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
 number of hits satisfying chosen parameters: 283224

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 100 summaries

```

```
Database :      PIR_73:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB	ID	Description
1	530	90.9	111	2	G30502	Ig kappa chain V r	
2	521	88.4	133	1	K4HUI1	Ig kappa chain pre	
3	515.5	88.4	114	1	K4HULN	Ig kappa chain V-I	
4	514.5	88.3	134	2	S34002	Ig kappa chain V r	
5	514.5	88.3	134	2	S49531	anti-5m antibody V	
6	513.5	88.1	113	2	S34003	Ig kappa chain V r	
7	512	87.8	112	2	S41393	Ig kappa chain V r	
8	511	87.7	138	2	S26040	Ig kappa chain pre	
9	510	87.5	112	2	S05970	Ig kappa chain V-J	
10	507.5	87.0	113	2	S30523	Ig kappa chain V r	
11	507.5	87.0	120	2	S51147	antibody light cha	
12	507	87.0	103	2	PH1054	Ig light chain V r	
13	502.5	86.2	113	2	S30520	Ig kappa chain V r	
14	500	85.8	118	2	PH0356	Ig kappa chain V r	
15	495.5	85.0	114	2	S44116	Ig kappa chain V-J	
16	495.5	85.0	240	2	S06084	Ig kappa chain pre	
17	494.5	84.8	114	2	S44119	Ig kappa chain V-J	
18	494.5	84.8	135	2	PL0014	Ig kappa chain pre	
19	492.5	84.5	134	1	K4HUI7	Ig kappa chain pre	
20	491	84.2	101	2	S26337	Ig light chain V r	
21	490.5	84.1	132	2	S46373	Ig kappa chain V-J	
22	489	83.9	132	2	S43103	Ig kappa chain V-J	
23	489	83.9	113	2	PT0407	Ig light chain V r	
24	486	83.4	138	2	A53261	Ig kappa chain pre	
25	485.5	83.3	124	2	S40364	Ig kappa chain - h	
26	484.5	83.1	134	2	PC1214	Ig kappa chain pre	
27	484	83.0	133	2	S50023	Ig kappa chain pre	
28	483.5	82.9	129	2	S06812	Ig kappa chain - h	
29	480.5	82.4	214	2	S48347	Ig kappa chain (Ma	


```
Query Match      88.3%; Score 514.5; DB 2; Length 113;
Best Local Similarity 88.5%; Pred. No. 1.2e-39;
Matches 100; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYYCHQYLS-SYTFGGGKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYYCHQYQYIPRTFGQGTKEIK 113
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
S49531
anti-Sm antibody VL chain (V kappa 4/J kappa 3) - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1995 #sequence_revision 12-May-1995 #text_change 21-Jan-2000
C:Accession: S49531
R:Amoudi, M.; Edwards, J.; Cairns, E.; Bell, D.
A:Submitted to the EMBL Data Library, October 1994
C:Description: Molecular characterization of natural human anti-Sm autoantibodies.
A:Reference number: S48797
A:Accession: S49531
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-134 <MAH>
A:Cross-references: EMBL:Z46347; NID:G560841; PIDN:CAA86466.1; PID:G560842
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:36-116/Domain: immunoglobulin homology <IMM>

Query Match      88.3%; Score 514.5; DB 2; Length 134;
Best Local Similarity 87.6%; Pred. No. 1.4e-39;
Matches 99; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYYCHQYLS-SYTFGGGKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYYCQYQYTAFTFGGTKEIK 133
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
S34003
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S34003
R:Marlette, X.; Tsapis, A.; Brouet, J.C.
A:Title: Nucleotide sequence analysis of the variable domains of four human monoclonal
Eur. J. Immunol. 23, 846-851, 1993
A:Reference number: S34001; MUID:93209281; PMID:7681398
A:Accession: S34003
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <MAH>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match      88.1%; Score 513.5; DB 2; Length 113;
Best Local Similarity 88.5%; Pred. No. 1.5e-39;
Matches 100; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYYCHQYLS-SYTFGGGKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYYCQYQYLTPTPTFGGTKEIK 113
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 7
S41393
Ig kappa chain V region (12.5H VL) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
C:Accession: S41393
R:Margritte, C.; Gilbert, D.; Brard, F.; Tron, F.
A:Submitted to the EMBL Data Library, January 1994
C:Description: Structural characterization of an (NZB X NZW)F1 mouse-derived IgM anti
A:Reference number: S41393
A:Accession: S41393
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112 <MAR>
A:Cross-references: EMBL:Z29536
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match      87.8%; Score 512; DB 2; Length 112;
Best Local Similarity 87.5%; Pred. No. 2e-39;
Matches 98; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYYCHQYLS-SYTFGGGKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 DSGVDPDRFSGSGGTDTLTISVQAEDLAVYYCKQSYNLTFTGGGKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
S26040
Ig kappa chain precursor - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 21-Jan-2000
C:Accession: S26040; S78098
R:Okamoto, M.; Honjo, T.
Nucleic Acids Res. 18, 1895, 1990
A:Title: Nucleotide sequences of the gene/cDNA coding for anti-murine erythrocyte aut
A:Reference number: S09216; MUID:90245589; PMID:2336368
A:Accession: S26040
A:Molecule type: DNA
A:Residues: 1-138 <OKA>
A:Cross-references: EMBL:X51742
A:Note: the authors translated the codon AGC for residue 107 as Thr and AGT for resid
submitted to the EMBL Data Library, February 1990
R:Okamoto, M.
A:Reference number: S78098
A:Accession: S78098
A:Molecule type: DNA
A:Residues: 1-87, 'W', 89-138 <OKW>
A:Cross-references: EMBL:X51742; NID:G52697; PIDN:CAA36032.1; PID:G52698
C:Genetics:
A:Introns: 22/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-138/Product: Ig kappa chain (fragment) #status predicted <MAT>
F:41-121/Domain: immunoglobulin homology <IMM>

Query Match      87.7%; Score 511; DB 2; Length 138;
Best Local Similarity 85.7%; Pred. No. 3.1e-39;
Matches 96; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNLAWYQKPGQSPKLLIYWASTR 85
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYYCHQYLS-SYTFGGGKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Query Match 86.2%; Score 502.5; DB 2; Length 113;
 Best Local Similarity 85.0%; Pred. No. 1.5e-38;
 Matches 96; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60
 Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60

Qy 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 112
 Db 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 113

RESULT 14
 PT0356
 Ig kappa chain V region (2B11.1) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jan-2000
 C:Accession: PT0356
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-240/Product: Ig kappa chain #status predicted <MAT>
 F:153-222/Domain: immunoglobulin homology <IMM>

Query Match 85.8%; Score 500; DB 2; Length 118;
 Best Local Similarity 85.6%; Pred. No. 2.6e-38;
 Matches 97; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60
 Db 4 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 63

Qy 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 112
 Db 64 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 115

RESULT 15
 S44119
 Ig kappa chain V-J region - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-May-2001
 C:Accession: S44119
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 85.0%; Score 495.5; DB 2; Length 114;
 Best Local Similarity 85.0%; Pred. No. 6.3e-38;
 Matches 96; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60
 Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60

Qy 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 112
 Db 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 113

RESULT 16
 S44119
 Ig kappa chain V-J region - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-May-2001
 C:Accession: S44119
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 84.8%; Score 494.5; DB 2; Length 114;
 Best Local Similarity 85.0%; Pred. No. 7.7e-38;
 Matches 96; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60
 Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60

Qy 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 112
 Db 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 113

RESULT 18
 PL0014
 Ig kappa chain precursor V region (F6-3) - mouse (fragment)

Qy 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 112
 Db 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 113

RESULT 16
 S06084
 Ig kappa chain precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jan-2000
 C:Accession: S06084
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-240/Product: Ig kappa chain #status predicted <MAT>
 F:153-222/Domain: immunoglobulin homology <IMM>

Query Match 85.0%; Score 495.5; DB 2; Length 240;
 Best Local Similarity 84.1%; Pred. No. 1.4e-37;
 Matches 95; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60
 Db 21 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 80

Qy 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 112
 Db 81 QSGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 133

RESULT 17
 S44119
 Ig kappa chain V-J region - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-May-2001
 C:Accession: S44119
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 84.8%; Score 494.5; DB 2; Length 114;
 Best Local Similarity 85.0%; Pred. No. 7.7e-38;
 Matches 96; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60
 Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQPKLLIYWASTR 60

Qy 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 112
 Db 61 ESGVDPDRFSGSGGTDTLTITSSVQAEADLAVYYCHQYLSS-YTFGGGKLEIK 113

RESULT 18
 PL0014
 Ig kappa chain precursor V region (F6-3) - mouse (fragment)

C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Jan-2000
C:Accession: PL0014
R:Cheng, H.L.; Sood, A.K.; Ward, R.E.; Kieber-Emmons, T.; Kohler, H.
Mol. Immunol. 25, 33-40, 1988
A:Title: Structural basis of stimulatory anti-idiotypic antibodies.
A:Reference number: PL0011; MUID:88142863; PMID:3125424
A:Accession: PL0014
A:Molecule type: mRNA
A:Residues: 1-145 <CHS>
A:Experimental source: cell line F6-3
A:Comment: This protein is an anti-idiotypic antibody that induces an anti-phosphorylcholine
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-134/Product: Ig heavy chain V region (4C11) #status predicted <MAT>
F:36-116/Domain: immunoglobulin homology <IMM>
F:44-60/Region: complementarity-determining 1
F:76-82/Region: complementarity-determining 2
F:115-121/Region: complementarity-determining 3
F:122-134/Region: complementarity-determining 4
F:143-144/Disulfide bonds: #status predicted <COR>
Query Match 84.8%; Score 494.5; DB 2; Length 145;
Best Local Similarity 84.2%; Pred. No. 9.9e-38;
Matches 96; Conservative 11; Mismatches 4; Indels 3; Gaps 2;
Qy 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQSPKLLIYWASTR 60
Db 21 DIVMSQSPSLAVSLGKVTMCKSSQSVLYSSNKNYLAWSYQKPGQSPKLLIYWASTR 80
Qy 61 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQYLSSY--TFGGGKLEIK 112
Db 81 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQY--DSYPLTFGSGTKLEMK 133
RESULT 19
K4HU17
Ig kappa chain precursor V-IV region (B17) - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 21-Jan-2000
C:Accession: A01905
R:Marsh, P.; Mills, F.; Gould, H.
Nucleic Acids Res. 13, 6531-6544, 1985
A:Title: Detection of a unique human V kappaIV germline gene by a cloned cDNA probe.
A:Reference number: A01905; MUID:86041854; PMID:2997713
A:Accession: A01905
A:Molecule type: mRNA
A:Residues: 1-134 <MAP>
Note: the sequence was determined from the differentiated gene
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger complexes.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-134/Product: Ig kappa chain V-IV region (B17) #status predicted <MAT>
F:21-43/Region: framework 1
F:36-116/Domain: immunoglobulin homology <IMM>
F:44-60/Region: complementarity-determining 1
F:61-75/Region: framework 2
F:76-82/Region: complementarity-determining 2
F:83-114/Region: framework 3
F:115-121/Region: complementarity-determining 3
F:122-134/Region: framework 4
F:143-144/Disulfide bonds: #status predicted
Query Match 84.5%; Score 492.5; DB 1; Length 134;
Best Local Similarity 85.0%; Pred. No. 1.4e-37;
Matches 96; Conservative 8; Mismatches 8; Indels 1; Gaps 1;
Qy 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQSPKLLIYWASTR 60
Db 21 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAWSYQKPGQPPKLLIYCASTR 80

Qy 61 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQYLS-SYTFGGGKLEIK 112
Db 81 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQYVNLPTFTFGGKVEIK 133
RESULT 20
S26337
Ig light chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 20-Jun-2000
C:Accession: S26337; S78449
R:Stark, S.E.; Caton, A.J.
J. Exp. Med. 174, 613-624, 1991
A:Title: Antibodies that are specific for a single amino acid interchange in a protein
A:Reference number: S26309; MUID:91341421; PMID:1908510
A:Accession: S26337
A:Molecule type: mRNA
A:Residues: 1-101 <STA>
A:Cross-references: EMBL:X59193
R:Caton, A.J.
submitted to the EMBL Data Library, April 1991
A:Reference number: S78447
A:Accession: S78449
A:Molecule type: mRNA
A:Residues: 1-60, 'T', 'S', '93-101 <CAT>
A:Cross-references: EMBL:X59193; NID:g52323; PIDN:CAA41903.1; PID:g1334067
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:8-88/Domain: immunoglobulin homology <IMM>
Query Match 84.2%; Score 491; DB 2; Length 101;
Best Local Similarity 93.0%; Pred. No. 1.4e-37;
Matches 93; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 10 SLAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQSPKLLIYWASTRESGVDRFS 69
Db 2 SLAVSAGEKVTMCKSSQSVLYSSNKNYLAWSYQKPGQSPKLLIYWASTRESGVDRFI 61
Qy 70 GSGGTDFLTITSSVQAEADLAVYYCHQYLSSTFTFGGKTL 109
Db 62 GSGGTDFLTITSSVQAEADLAVYYCHQYLSSTFTFGGKTL 101
RESULT 21
S46373
Ig kappa chain V-J region (T24-9) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 27-Jan-1995 #sequence_revision 01-Sep-1995 #text_change 21-Jan-2000
C:Accession: S46373; S38647
R:Bensimon, C.; Chastagner, P.; Zouali, M.
EMBO J. 13, 2951-2962, 1994
A:Title: Human lupus anti-DNA autoantibodies undergo essentially primary V(chI) gene
A:Reference number: S46369; MUID:94313975; PMID:8039491
A:Accession: S46373
A:Molecule type: mRNA
A:Residues: 1-132 <BEN>
A:Cross-references: EMBL:Z27174; NID:g415963; PIDN:CAA81698.1; PID:g415964
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:34-114/Domain: immunoglobulin homology <IMM>
Query Match 84.1%; Score 490.5; DB 2; Length 132;
Best Local Similarity 84.1%; Pred. No. 2.1e-37;
Matches 95; Conservative 9; Mismatches 8; Indels 1; Gaps 1;
Qy 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGQSPKLLIYWASTR 60
Db 19 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAWSYQKPGQPPKLLIHWASSR 78
Qy 61 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQYLS-SYTFGGGKLEIK 112
Db 79 ESGVDPDRFSGSGGTDFLTITSSVQAEADLAVYYCHQYVNLPTFTFGGKVDIK 131

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:10 ; Search time 14.4965 Seconds
(without alignments)
1591.926 Million cell updates/sec

Title: US-10-056-052A-18

Perfect score: 583

Sequence: 1 DIVMTQSPDLSAVSLGERTV.....YCHQYLSSTYFGGKTLEIK 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

SPREMBL_21:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mhc:.*
8: sp_organelle:.*
9: sp_phage:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_virus:.*
16: sp_bacteriophage:.*
17: sp_archaea:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	432.5	74.2	107	11 Q9ERZ9	Q9ERZ9 mus musculus
2	391	67.1	109	4 Q9UL78	Q9UL78 homo sapien
3	389	66.7	107	4 Q96SA9	Q96SA9 homo sapien
4	383	65.7	238	11 Q99M37	Q99M37 mus musculus
5	378.5	64.9	108	4 Q9UL70	Q9UL70 homo sapien
6	378	64.8	109	4 Q9UL85	Q9UL85 homo sapien
7	377	64.7	238	11 Q9VC16	Q9VC16 mus musculus
8	376.5	64.6	108	4 Q9UL77	Q9UL77 homo sapien
9	376.5	64.6	111	11 Q920E9	Q920E9 mus musculus
10	376	64.5	107	4 Q9UL81	Q9UL81 homo sapien
11	372.5	63.9	108	4 Q9UL83	Q9UL83 homo sapien
12	371.5	63.7	108	4 Q9UL79	Q9UL79 homo sapien
13	367	63.0	109	4 Q9UL86	Q9UL86 homo sapien
14	366	62.8	239	11 Q8VC55	Q8VC55 mus musculus
15	363.5	62.3	99	11 Q9JL74	Q9JL74 mus musculus
16	363.5	62.3	108	11 Q8VIJ0	Q8VIJ0 mus musculus

17	353.5	60.6	114	4 Q9UL80	Q9UL80 homo sapien
18	350	60.0	239	4 Q8TCD0	Q8TCD0 homo sapien
19	349	59.9	104	11 Q9JL82	Q9JL82 mus musculus
20	344.5	59.1	233	11 Q9LWS9	Q9LWS9 mus musculus
21	341.5	58.6	214	11 Q9RIA5	Q9RIA5 mus musculus
22	335.5	57.5	234	11 Q8R062	Q8R062 mus musculus
23	335.5	57.5	298	11 Q9QYF0	Q9QYF0 mus musculus
24	333.5	57.2	109	11 Q920E6	Q920E6 mus musculus
25	333	57.1	134	11 Q8VDD0	Q8VDD0 mus musculus
26	321.5	55.1	103	11 Q9JL80	Q9JL80 mus musculus
27	320.5	55.0	107	11 Q9JL84	Q9JL84 mus musculus
28	320.5	55.0	116	4 Q96PF6	Q96PF6 homo sapien
29	320.5	55.0	234	11 Q9LWF8	Q9LWF8 mus musculus
30	315.5	54.1	101	11 Q9JL78	Q9JL78 mus musculus
31	313	53.7	106	5 Q9U410	Q9U410 schistosoma
32	312.5	53.6	234	11 Q8VCP0	Q8VCP0 mus musculus
33	311.5	53.4	97	11 Q9JL76	Q9JL76 mus musculus
34	311.5	53.4	127	11 Q925S9	Q925S9 mus musculus
35	302	51.8	241	11 Q92IA6	Q92IA6 mus musculus
36	301	51.6	235	11 Q9LW12	Q9LW12 mus musculus
37	290	49.7	130	4 Q9NP29	Q9NP29 homo sapien
38	287.5	49.3	109	6 Q9NOW5	Q9NOW5 oryctolagus
39	265.5	45.5	234	11 Q8R028	Q8R028 mus musculus
40	263	45.1	237	4 Q8WUK4	Q8WUK4 homo sapien
41	259	44.4	233	4 Q8TBC9	Q8TBC9 homo sapien
42	259	44.4	237	4 Q8WTU6	Q8WTU6 homo sapien
43	251	43.1	107	4 Q9NSD6	Q9NSD6 homo sapien
44	244	41.9	110	4 Q8TE63	Q8TE63 homo sapien
45	236.5	40.6	107	4 Q9UL82	Q9UL82 homo sapien
46	232	39.8	218	11 Q925S1	Q925S1 mus musculus
47	231.5	39.7	108	4 Q96SB0	Q96SB0 homo sapien
48	231.5	39.7	236	4 Q96E61	Q96E61 homo sapien
49	229.5	39.4	112	4 Q96JD1	Q96JD1 homo sapien
50	229	39.3	116	4 Q96JD0	Q96JD0 homo sapien
51	221	37.9	235	11 Q99M11	Q99M11 mus musculus
52	216	37.0	112	4 Q96JD2	Q96JD2 homo sapien
53	199	34.1	109	11 Q9ET13	Q9ET13 mus musculus
54	192.5	33.0	240	4 Q8WUK3	Q8WUK3 homo sapien
55	191.5	32.8	129	11 Q8VDE2	Q8VDE2 mus musculus
56	175	30.0	233	4 Q96I69	Q96I69 homo sapien
57	168.5	28.9	132	4 Q8TBD0	Q8TBD0 homo sapien
58	141	24.2	135	4 Q9H5Z4	Q9H5Z4 homo sapien
59	135.5	23.2	93	4 Q9UL76	Q9UL76 homo sapien
60	133	22.8	123	11 Q61243	Q61243 mus musculus
61	124	21.3	143	11 Q924Q0	Q924Q0 mus musculus
62	120	20.6	206	13 Q8UWK3	Q8UWK3 ictalurus p
63	120	20.6	223	13 Q8UWK8	Q8UWK8 ictalurus p
64	118.5	20.3	144	11 Q924P5	Q924P5 mus musculus
65	117	20.1	337	13 Q9IB02	Q9IB02 spherooides
66	115.5	19.8	117	13 Q8UV75	Q8UV75 brachydanio
67	115.5	19.8	325	13 Q8UWL3	Q8UWL3 ictalurus p
68	115	19.7	195	13 Q8UV39	Q8UV39 brachydanio
69	114	19.6	340	13 Q9IAZ0	Q9IAZ0 spherooides
70	113.5	19.5	142	11 Q924Q2	Q924Q2 mus musculus
71	113.5	19.5	320	13 Q9IAZ9	Q9IAZ9 spherooides
72	113	19.4	185	13 Q8UWK4	Q8UWK4 ictalurus p
73	113	19.4	280	13 Q8UWL1	Q8UWL1 ictalurus p
74	113	19.4	280	13 Q8UWK1	Q8UWK1 ictalurus p
75	112	19.2	206	13 Q8UWK5	Q8UWK5 ictalurus p
76	111.5	19.1	235	6 Q9XSM6	Q9XSM6 salmifri sci
77	111.5	19.1	484	11 Q8VEA0	Q8VEA0 mus musculus
78	111	19.0	145	11 Q924Q6	Q924Q6 mus musculus
79	111	19.0	340	13 Q9IAZ6	Q9IAZ6 spherooides
80	110.5	19.0	140	11 Q924P8	Q924P8 mus musculus
81	110.5	19.0	202	13 Q8UWK6	Q8UWK6 ictalurus p
82	110.5	19.0	227	13 Q8UWK9	Q8UWK9 ictalurus p
83	110	18.9	145	11 Q924Q7	Q924Q7 mus musculus
84	110	18.9	222	13 Q8UV38	Q8UV38 brachydanio
85	110	18.9	246	13 Q8UVA4	Q8UVA4 brachydanio
86	110	18.9	336	13 Q8UV73	Q8UV73 brachydanio
87	109.5	18.8	138	13 Q9W6Z6	Q9W6Z6 ginglymosto
88	109	18.7	110	11 Q9JL77	Q9JL77 mus musculus
89	108.5	18.6	140	11 Q924R2	Q924R2 mus musculus

90 108 18.5 143 11 Q924R0 Q924r0 mus musculus
 91 108 18.5 143 11 Q924P9 Q924p9 mus musculus
 92 108 18.5 224 13 Q8UV32 Q8uv32 brachydanio
 93 108 18.5 260 13 Q8UV96 Q8uv96 brachydanio
 94 108 18.5 275 13 Q8UV91 Q8uv91 brachydanio
 95 107.5 18.4 308 13 Q8UUG3 Q8uug3 ictalurus p
 96 107 18.4 143 11 Q91V67 Q91v67 mus musculus
 97 107 18.4 145 11 Q924R1 Q924r1 mus musculus
 98 106.5 18.3 146 11 Q924Q3 Q924q3 mus musculus
 99 106.5 18.3 168 4 Q9U056 Q9u56 homo sapien
 100 106.5 18.3 215 13 Q8UWK7 Q8uwk7 ictalurus p

ALIGNMENTS

RESULT 1

Q9ER29 PRELIMINARY; PRT; 107 AA.
 ID Q9ER29
 Q9ER29;
 01-MAR-2001 (TReMBLrel. 16, Created)
 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Anti human TNF-alpha light chain variable region (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN
 RP SEQUENCE FROM N.A.
 RA Chen P., Deng J.B., Wang Z.L., Han H., Su C.Z.;
 RT "Cloning and sequencing of the light chain fragment of variable region
 RT genes of an anti-hnrf-a monoclonal antibody.";
 RL J. Cell. Mol. Immunol. 12:21-26(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;
 RT "Construction and sequencing of the single-chain antibody gene of a
 RT human TNF-alpha specific monoclonal antibody.";
 RL Ti 4 Chun 1 Ta Hsueh Hsueh Pao 19:373-376(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF262753; AAG23804.1;
 DR HSSP; P80362; IWTL.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00406; IGV; 1.
 FT NON_TER 107
 SQ SEQUENCE 107 AA; 11784 MW; 2B15EEA6604A26C3 CRC64;

Query Match 74.2%; Score 432.5; DB 11; Length 107;
 Best Local Similarity 77.6%; Pred. No. 1.8e-38;
 Matches 83; Conservative 13; Mismatches 10; Indels 1; Gaps 1;
 QY 4 MTQSPSLAVSLGERTVMTCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTRESG 63
 Db 1 MTQSPSLAVSLGERTVMTCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTRESG 63
 QY 64 VPDPRFSGSGGTDTLTITSSVQEDLAVYYCHO-YLSSVTEFGGGTKL 109
 Db 61 VPDPRFSGSGGTDTLTITSSVQEDLADYFCQOHYRTPFTFGSGTKL 107

RESULT 2

Q9UL78 PRELIMINARY; PRT; 109 AA.
 ID Q9UL78
 AC Q9UL78
 01-MAY-2000 (TReMBLrel. 13, Created)
 QY 1 DIVMTQSPDLSAVSLGERTVMTCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
 Db 1 DIVMTQSPDLSAVSLGERTVMTCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60

01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Myosin-reactive immunoglobulin light chain variable region
 DE (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus.";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035036; AAD56272.1;
 DR HSSP; P80362; IWTL.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR NON_TER 1
 FT NON_TER 109
 SQ SEQUENCE 109 AA; 11646 MW; 5F675C52EC7EE197 CRC64;
 Query Match 67.1%; Score 391; DB 4; Length 109;
 Best Local Similarity 67.3%; Pred. No. 4.7e-34;
 Matches 76; Conservative 18; Mismatches 13; Indels 6; Gaps 2;
 QY 1 DIVMTQSPDLSAVSLGERTVMTCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
 Db 1 EIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAWYQKPGQAPRLIYGA 55
 QY 61 ESGVDPFRSGSGGTDTLTITSSVQEDLAVYYCHOYLSS-YTFGGGKLEIK 112
 Db 56 ATGIPDRFSGSGGTDTLTITSRLEPEDCAVYCCQYGGSSPLTFGGGKVEIK 108
 RESULT 3
 ID Q96SA9 PRELIMINARY; PRT; 107 AA.
 AC Q96SA9;
 01-DEC-2001 (TReMBLrel. 19, Created)
 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DE Anti-streptococcal/anti-myosin immunoglobulin kappa light chain
 DE variable region (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98375893; PubMed=9712075;
 RA Adderson E.E., Shikhan A.R., Ward K.E., Cunningham M.W.;
 RT "Molecular analysis of polyreactive monoclonal antibodies from
 RT rheumatic carditis: human anti-N-acetylglucosamine/anti-myosin
 RT antibody v region genes.";
 RL J. Immunol. 161:2020-2031(1998).
 DR EMBL; U96396; AAB68785.1;
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 1.
 DR NON_TER 1
 FT NON_TER 107
 SQ SEQUENCE 107 AA; 11520 MW; 4BB43E9C5B577F16 CRC64;
 Query Match 66.7%; Score 389; DB 4; Length 107;
 Best Local Similarity 67.0%; Pred. No. 7.4e-34;
 Matches 75; Conservative 15; Mismatches 16; Indels 6; Gaps 1;
 QY 1 DIVMTQSPDLSAVSLGERTVMTCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
 Db 1 DIVMTQSPDLSAVSLGERTVMTCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60

```
Db 1 DIQWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 54
Qy 61 ESGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLSVYTFGGGTGKLEIK 112
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 55 QSGVPSRFSGSGCTDFTLTISSLPQEDFATYYCQSYSTLTFTGGGTGKLEIK 106
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 4
Q99M37 PRELIMINARY; PRT; 238 AA.
AC Q99M37;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 26.3 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
Straussberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020335; AA02035.1; -.
DR HSSP; P01679; 2FBJ.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig.cl.
DR InterPro; IPR003600; Ig.like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 1.
DR SMART; SM00406; IGv; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; Ig_MHC; UNKNOWN_1.
DR Hypothetical protein.
KW PROSITE.
SQ SEQUENCE 238 AA; 26344 MW; FB2B06A0B801330A CRC64;

Query Match 65.78; Score 383; DB 11; Length 238;
Best Local Similarity 69.08; Pred. No. 9e-33;
Matches 78; Conservative 12; Mismatches 21; Indels 2; Gaps 2;

Qy 1 DIVWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 20 DVWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 78
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

61 ESGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 112
79 FSGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 131

RESULT 5
Q9UL70 PRELIMINARY; PRT; 108 AA.
AC Q9UL70;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=98277139; PubMed=9614934;
Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
Young D.C.;
"Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).

Qy 1 DIVWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 20 DVWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 78
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

61 ESGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 112
79 FSGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 131

RESULT 6
Q9UL85 PRELIMINARY; PRT; 109 AA.
AC Q9UL85;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin kappa chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=98277139; PubMed=9614934;
Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
Young D.C.;
"Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).

Qy 1 DIVWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 20 DVWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 78
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

61 ESGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 112
79 FSGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 131

RESULT 7
Q8VCI6 PRELIMINARY; PRT; 238 AA.
AC Q8VCI6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
```

```
DR EMBL; AF035044; AAD56280.1; -.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGv; 1.
FT NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 11633 MW; B7BEDC3E41FCCA37 CRC64;

Query Match 64.98; Score 378.5; DB 4; Length 108;
Best Local Similarity 67.38; Pred. No. 9.8e-33;
Matches 76; Conservative 13; Mismatches 17; Indels 7; Gaps 2;

Qy 1 DIVWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 DIQWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 54
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

61 ESGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 112
79 FSGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 107

RESULT 6
Q9UL85 PRELIMINARY; PRT; 109 AA.
AC Q9UL85;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin kappa chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=98277139; PubMed=9614934;
Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
Young D.C.;
"Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).

Qy 1 DIVWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 DIQWTSPPSLASVSGDRVTITCRASQSI-----SSYLNMYOQKPGKAPKLLIYAASL 54
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

61 ESGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 112
79 FSGVDPDRFSGSGCTDFTLTISSVQAEADLAVYYCHQVLS-SYTFGGGTGKLEIK 108

RESULT 7
Q8VCI6 PRELIMINARY; PRT; 238 AA.
AC Q8VCI6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
```

01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 26.2 kDa protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC019760; AAH19760.1; -;
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003597; IG.cl.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; Ig; 2.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00407; IGcl; 1.
 DR SMART; SM00406; IGV; 1.
 PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 Hypothetical protein
 SQ SEQUENCE 238 AA; 26224 MW; 35EC08E3DE5414AD CRC64;
 Query Match 64.7%; Score 377; DB 11; Length 238;
 Best Local Similarity 67.8%; Pred. No. 3.9e-32;
 Matches 78; Conservative 15; Mismatches 16; Indels 6; Gaps 3;
 QY 1 DIVVTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
 Db DVVNTQTPSLVSLGDSQASISCRSSQSLVH-SNGNTYLHWYKQPGQSPKLLIYKVSNR 78
 QY 61 ESGVDPDRFSGSGSGTDFTLTLSVQAEDLAVYYCHQYVSSY--TFGGGKLEIK 112
 Db 79 FSGVDPDRFSGSGSGTDFTLTKSRVEADLGVIYFCQ--STHVPPTFGGKLEIK 131
 RESULT 8
 Q9UL77 PRELIMINARY; PRT; 108 AA.
 ID Q9UL77
 AC Q9UL77
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Myosin-reactive immunoglobulin light chain variable region
 DE (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035037; AAD56273.1; -;
 DR HSP; P01607; IREI.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR SMART; SM00409; IG; 1.
 FT NON_TER 108
 FT SEQUENCE 108 AA; 11738 MW; C06681716C4D16F3 CRC64;
 Query Match 64.6%; Score 376.5; DB 4; Length 108;
 Best Local Similarity 65.5%; Pred. No. 1.6e-32;
 Matches 74; Conservative 16; Mismatches 16; Indels 7; Gaps 2;
 QY 1 DIVVTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60

Db 1 DIQMTQSPSSLSASVSGDRVTITCRASQSI-----SSYLNWYQKPGKAPNLLIYAASL 54
 QY 61 ESGVDPDRFSGSGSGTDFTLTLSVQAEDLAVYYCHQ-YLSSYTFGGGKLEIK 112
 Db 55 QSGVPSRFSGSGSGTDFTLTLSVQAEDLAVYYCHQYVSSYTFGGGKLEIK 107
 RESULT 9
 Q920E9 PRELIMINARY; PRT; 111 AA.
 ID Q920E9
 AC Q920E9
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Pterin-mimicking anti-idiotope kappa chain variable region
 DE (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Atkin J.D., Iape A., Jennings I.G., Horaitis O., Cotton R.G.H.;
 RT "Definition of the Idiotope of Pterin-Mimicking Antibodies Expressed in Mammalian Cells."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF079335; AAL09419.1; -;
 DR InterPro; IPR003006; IG_MHC.
 DR Pfam; PF00047; Ig; 1.
 FT NON_TER 111
 FT SEQUENCE 111 AA; 12046 MW; 1E46988AA6858526 CRC64;
 Query Match 64.6%; Score 376.5; DB 11; Length 111;
 Best Local Similarity 68.1%; Pred. No. 1.7e-32;
 Matches 77; Conservative 12; Mismatches 21; Indels 3; Gaps 2;
 QY 1 DIVVTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
 Db 1 DIVLTQSPASVSLGQRATISCRASKSV--STGYSYTHWYQKPGQPPKLLIYASNL 58
 QY 61 ESGVDPDRFSGSGSGTDFTLTLSVQAEDLAVYYCHQYVSSYTFGGGKLEIK 112
 Db 59 ESGVDPDRFSGSGSGTDFTLTNHPVEEDAATYQCQSRELPTFTFGGKLEIK 111
 RESULT 10
 Q9UL81 PRELIMINARY; PRT; 107 AA.
 ID Q9UL81
 AC Q9UL81
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Myosin-reactive immunoglobulin light chain variable region
 DE (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035033; AAD56269.1; -;
 DR HSP; P01607; IREI.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; Ig; 1.


```
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;

Query Match 64.5%; Score 376.; DB 4; Length 107;
Best Local Similarity 64.3%; Pred. No. 1.8e-32;
Matches 72; Conservative 16; Mismatches 18; Indels 6; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWOQKPGQSPKLLIYWASTR 60
Db 1 DIQMTQSPSLASVGDRTVITCRASQSI-----SNLYNWYQKPGKAPNLLIYRASSL 54
QY 61 ESGVPRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLS-SYTFGGTKLEIK 112
Db 55 QSGVPRFSGSGGTDTFTLTISGLQAEAFATYCCQSALTEGPGTKVDIR 106

RESULT: 11
Q9UL83 PRELIMINARY; PRT; 108 AA.
Q9UL83;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=98277139; PubMed=9614934;
Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
Young D.C.;
RA "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035035; AAD56267.1; -
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11834 MW; 9FC5A92EBA96E6A CRC64;

Query Match 63.9%; Score 372.5; DB 4; Length 108;
Best Local Similarity 67.3%; Pred. No. 4.3e-32;
Matches 76; Conservative 17; Mismatches 13; Indels 7; Gaps 3;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWOQKPGQSPKLLIYWASTR 60
Db 1 EIVMTQSPATLSLPPGERATLSCRASQSV--SSN----LAWYQKPGQAPRLIYCASTR 54
QY 61 ESGVPRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLS-SYTFGGTKLEIK 112
Db 55 ATGIPARFSGSGGTDTFTLTISLQAEAFATYCCQHNWNPFTFGPGTKVDIK 107

RESULT 12
Q9UL79 PRELIMINARY; PRT; 108 AA.
Q9UL79;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=98277139; PubMed=9614934;
Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
Young D.C.;
RA "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035035; AAD56271.1; -
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11787 MW; DB5845F19724FB4E CRC64;

Query Match 63.7%; Score 371.5; DB 4; Length 108;
Best Local Similarity 65.5%; Pred. No. 5.5e-32;
Matches 74; Conservative 14; Mismatches 18; Indels 7; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWOQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPSLASVGDRTVITCRMSQSI-----SSYLAWOQKPGKAPPELLIYAATL 54
QY 61 ESGVPRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLS-SYTFGGTKLEIK 112
Db 55 QSGVPRFSGSGGTDTFTLTISLQAEAFATYCCQYSPFTFGPGTKVEIK 107

RESULT 13
Q9UL86 PRELIMINARY; PRT; 109 AA.
Q9UL86;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin kappa chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=98277139; PubMed=9614934;
Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
Young D.C.;
RA "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035028; AAD56264.1; -
DR HSSP; P80362; IWL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 11928 MW; 243325F72C7DAC83 CRC64;

Query Match 63.0%; Score 367.; DB 4; Length 109;
Best Local Similarity 62.8%; Pred. No. 1.7e-31;
Matches 71; Conservative 20; Mismatches 16; Indels 6; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWOQKPGQSPKLLIYWASTR 60
Db 1 EIVTQSPGTSLSPGERATLSCRASQSV-----SSSYLAWOQKPGQAPRLIYGTSSR 55
QY 61 ESGVPRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLS-SYTFGGTKLEIK 112
```


OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035034; AAD36270.1; -
DR HSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
FT NON_TER 1 114
SEQUENCE 114 AA; 12775 MW; -070E31E210D1CB01 CRC64;

Query Match 60.6%; Score 353.5; DB 4; Length 114;
Best Local Similarity 62.3%; Pred. No. 4.8e-30;
Matches 71; Conservative 17; Mismatches 23; Indels 3; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
DB 1 DVMTQSPSLPVLTRQPAISICRSPVSDG-NTYLNWFQRPQSPRLIYKVSNR 59
:
QY 61 ESGVDPDRSGSGGDTFTLTSSVQAEADLAVYCHQ--YLSYTFGGGKLEIK 112
DB 60 DSGVDPDRSGSGGDTFTLKISRVEADGVVYCHQGHWPPTFGQCKVEIK 113
:

RESULT 18
Q8TCD0 PRELIMINARY; PRT; 239 AA.
AC Q8TCD0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 26.2 kDa protein.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022362; AAH22362.1; -
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26234 MW; FAGEDC3A3B03871D CRC64;

Query Match 60.0%; Score 350; DB 4; Length 239;
Best Local Similarity 61.9%; Pred. No. 2.9e-29;
Matches 70; Conservative 18; Mismatches 23; Indels 2; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
DB 21 DVMTQSPSLPVLTRQPAISICRSTQSLVSDG-NTYLNWFQRPQSPRLIYKVSNR 79
:
QY 61 ESGVDPDRSGSGGDTFTLTSSVQAEADLAVYCHQYLS-SYTFGGGKLEIK 112
DB 80 DSGVDPDRSGSGGDTFTLKISRVEADGVYFCMQGHWPSTFGQCKLEIK 132
:

RESULT 19
Q9JL82 PRELIMINARY; PRT; 104 AA.
ID Q9JL82
AC Q9JL82;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Anti-myosin immunoglobulin light chain variable region
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAUB/C;
RX MEDLINE=20448942; PubMed=10992488;
RA Malkiel S., Liao L., Cunningham M.W., Diamond B.;
RT "T-Cell-dependent antibody response to the dominant epitope of
RT streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive
RT with cardiac myosin";
RL Infect. Immun. 68:5803-5808(2000).
DR EMBL; AF206024; AAF69322.1; -
DR HSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
FT NON_TER 1 104
FT NON_TER 104 104
SQ SEQUENCE 104 AA; 11360 MW; 5DA8BBD5F0AA1AE CRC64;

Query Match 59.9%; Score 349; DB 11; Length 104;
Best Local Similarity 68.3%; Pred. No. 1.3e-29;
Matches 71; Conservative 12; Mismatches 19; Indels 2; Gaps 2;

QY 10 SLAVTSQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTRSGVDPDRFS 69
DB 2 SLPVSLGQASISCRSSQSLVH-TNGNTYLHWYLOKPGQSPKLLIYKVSNRSGVDPDRFS 60
:
QY 70 GSGGDTFTLTSSVQAEADLAVYCHQYLS-SYTFGGGKLEIK 112
DB 61 GSGGDTFTLKISRVEADGVYFCSTQTHVPYTFGGGKLEIK 104
:

RESULT 20
Q9IWS9 PRELIMINARY; PRT; 233 AA.
ID Q9IWS9;
AC Q9IWS9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 25.8 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013496; AAH13496.1; -
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
FT NON_TER 1 233
SQ SEQUENCE 233 AA; 25781 MW; BIC184DA149A16EB CRC64;

Query Match 59.1%; Score 344.5; DB 11; Length 233;
Best Local Similarity 60.2%; Pred. No. 1.1e-28;
Matches 68; Conservative 16; Mismatches 22; Indels 7; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
DB 20 DIQMTQTSSLSASLGDRVTISCSGSGI-----ANYLNWYQKPDGTGKLLIYSSSL 73
:

```
Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEADLAVYYCHQY-LSSYFGGGTKLEIK 112
|||||
Db 74 HSGVPSRFSGSGGTDTFTLTSSVQAEADLAVYYCHQY-LSSYFGGGTKLEIK 126
|||||

RESULT 21
Q9RIA5 PRELIMINARY; PRT; 214 AA.
ID Q9RIA5;
AC Q9RIA5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Kappa light chain of Mab7 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scfv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152371; AAD40242.1; -
DR HSP; P01679; 2FBJ.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00406; IG; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 214
SQ SEQUENCE 214 AA; 23922 MW; 52BA205FDE995E2A CRC64;

Query Match
Best Local Similarity 58.6%; Score 341.5; DB 11; Length 214;
Matches 67; Conservative 17; Mismatches 22; Indels 7; Gaps 2;

Qy 1 DIVWTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSKLLIYWASTR 60
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 DIQLTQSPSSAYSLGERVTITCKASQDI-----NSYLSWFOKPGKSPKTLIYRANRL 54
|||||

Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEADLAVYYCHQYLS-SYTFGGGTKLEIK 112
|||||
Db 55 VDGVPFRFSGSGGTDTFTLTSSVQAEADLAVYYCHQYLS-SYTFGGGTKLEIK 107
|||||

RESULT 22
Q8R062 PRELIMINARY; PRT; 234 AA.
ID Q8R062;
AC Q8R062;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical 25.9 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027418; AAH27418.1; -
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 234 AA; 25857 MW; 4EB08C81426AEB1 CRC64;

Query Match
Best Local Similarity 57.5%; Score 335.5; DB 11; Length 234;
```

```
Best Local Similarity 58.4%; Pred. No. 9.9e-28;
Matches 66; Conservative 17; Mismatches 23; Indels 7; Gaps 2;

Qy 1 DIVWTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSKLLIYWASTR 60
|||||
Db 21 DIQLTQSPSSAYSLGERVTITCKASQDI-----SNYLSWFOKPGKSPKTLIYRANRL 74
|||||

Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEADLAVYYCHQYLS-SYTFGGGTKLEIK 112
|||||
Db 75 HSGVPSRFSGSGGTDTFTLTSSVQAEADLAVYYCHQYLS-SYTFGGGTKLEIK 127
|||||

RESULT 23
Q9QYF0 PRELIMINARY; PRT; 298 AA.
ID Q9QYF0;
AC Q9QYF0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CN 8 scfv.
GN CN 8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-SPLEEN;
RX MEDLINE=20183931; PubMed=10706631;
RA Shinohara N., Demura T., Fukuda H.;
RT "Isolation of a vascular cell wall-specific monoclonal antibody
RT recognizing a cell polarity by using a phage display subtraction
RT method.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).
DR EMBL; AB036341; BAA88633.1; -
DR HSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00406; IG; 2.
DR SMART; SM00410; IG_Like; 2.
SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match
Best Local Similarity 57.5%; Score 335.5; DB 11; Length 298;
Matches 67; Conservative 15; Mismatches 24; Indels 7; Gaps 2;

Qy 1 DIVWTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSKLLIYWASTR 60
|||||
Db 173 DIQLTQSPSSAYSLGERVTITCKASQDI-----SGNIHNYLAWYQKPGKSPKTLIYRANRL 226
|||||

Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEADLAVYYCHQYLS-SYTFGGGTKLEIK 112
|||||
Db 227 ADGVPSRFSGSGGTDTFTLTSSVQAEADLAVYYCHQYLS-SYTFGGGTKLEIK 279
|||||

RESULT 24
Q920E6 PRELIMINARY; PRT; 109 AA.
ID Q920E6;
AC Q920E6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Pterin-mimicking anti-idiotope kappa chain variable region
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Atkin J.D., Iape A., Jennings I.G., Horaitis O., Cotton R.G.H.;
RT "Definition of the Idiotope of Pterin-Mimicking Antibodies Expressed
RT in Mammalian Cells.";
```

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF307938; AAL09422.1; -

DR InterPro: IPR003006; Ig_MHC.

DR Pfam: PF00047; Ig_1.

FT NON_TER 1

FT NON_TER 109

SQ SEQUENCE 109 AA; 11943 MW; DAD3F98E05DD1501 CRC64;

Query Match 57.2%; Score 333.5; DB 11; Length 109;

Best Local Similarity 59.3%; Pred. No. 6e-28;

Matches 67; Conservative 15; Mismatches 24; Indels 7; Gaps 2;

QY 1 DIVTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGOSP KLLIYWASTR 60

DB 1 DIQMTSPASLSASVGTITTCRA-----SGNIHNYLAWYQKQKSPQLLYVNAKTL 54

QY 61 ESGVPDRFSGSGGTDFTLTSSVQAEDLAVVYCHQYLSS-YTGGGKLEIK 112

DB 55 ADGVPSRFSGSGGSGTQYSLKINSIQPEDFGSYCOHFWSPTWTFGGGKLEIK 107

SUBMIT 25

Q8VDDO

ID Q8VDDO PRELIMINARY; PRT: 134 AA.

AC Q8VDDO:

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Anti-MOG z12 variable light chain (Fragment).

GN ANTI-MOG KAPPA.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALE/C;

RA Chernajovsky Y.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALE/C;

RA Sembl P.;

RT "Targeting T cells to the CNS.";

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ416331; CAC94866.1; -

DR InterPro: IPR003599; Ig.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003596; Ig_V.

DR Pfam: PF00047; Ig_1.

DR SMART: SM00409; IG; 1.

DR SMART: SM00406; IGV; 1.

FT NON_TER 134

FT NON_TER 134

SQ SEQUENCE 134 AA; 14525 MW; CDFE8E2236E2D0CF CRC64;

Query Match 57.1%; Score 333; DB 11; Length 134;

Best Local Similarity 61.9%; Pred. No. 8.9e-28;

Matches 70; Conservative 12; Mismatches 21; Indels 10; Gaps 3;

QY 2 IVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGOSP KLLIYWASTR 61

DB 24 IVLTQSPAIMASPGKVTMTCSASSI-----SYNHWYQKPGTSPKRMIYDTSKLA 76

QY 62 SGVPDRFSGSGGTDFTLTSSVQAEDLAVVYCHQYLSS--TFGGGKLEIK 112

DB 77 SGVPARFSGSGSGTYSYSLTSSMEADAATYCHO-RSSYPWTFGGGKLEIK 128

Search completed: June 23, 2003, 14:03:30

Job time : 15.4965 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 4.17021 Seconds
(without alignments)
1113.936 Million cell updates/sec

Title: US-10-056-052a-18
Perfect score: 583
Sequence: 1 DIVMTQSPDLSAVSLGERTV.....YCHQYLSSYTFGGGTKLEIK 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	521	89.4	133	1	KV4B_HUMAN	P06313 homo sapien
2	520.5	89.3	114	1	KV4A_HUMAN	P01625 homo sapien
3	505.5	86.7	134	1	KV4C_HUMAN	P06314 homo sapien
4	479	82.2	121	1	KV40_HUMAN	P06312 homo sapien
5	388.5	66.6	136	1	KV5B_MOUSE	P01634 mus musculus
6	387	66.4	110	1	KV3P_MOUSE	P01668 mus musculus
7	384.5	66.0	111	1	KV3L_MOUSE	P01664 mus musculus
8	381	65.4	129	1	KV3L_HUMAN	P18135 homo sapien
9	381	65.4	129	1	KV3M_HUMAN	P18136 homo sapien
10	379.5	65.1	149	1	KV5A_MOUSE	P01633 mus musculus
11	377	64.7	109	1	KV3B_HUMAN	P01620 homo sapien
12	376	64.5	109	1	KV3D_HUMAN	P01622 homo sapien
13	375.5	64.4	114	1	KV1A_MOUSE	P01632 mus musculus
14	374.5	64.2	108	1	KV1M_HUMAN	P01605 homo sapien
15	374	64.2	109	1	KV3E_HUMAN	P01623 homo sapien
16	373.5	64.1	111	1	KV3M_MOUSE	P01665 mus musculus
17	373.5	64.1	111	1	KV3Q_MOUSE	P01667 mus musculus
18	372.5	63.9	111	1	KV3H_MOUSE	P01660 mus musculus
19	372	63.8	129	1	KV3H_HUMAN	P04207 homo sapien
20	367.5	63.0	111	1	KV3N_MOUSE	P01666 mus musculus
21	367.5	63.0	111	1	KV3R_MOUSE	P01670 mus musculus
22	367.5	63.0	129	1	KV1W_HUMAN	P04311 homo sapien
23	367	63.0	108	1	KV3A_HUMAN	P01619 homo sapien
24	367	63.0	109	1	KV3F_HUMAN	P01624 homo sapien
25	366.5	62.9	111	1	KV3J_MOUSE	P01662 mus musculus
26	366	62.8	133	1	KV2F_HUMAN	P06310 homo sapien
27	365.5	62.7	111	1	KV3Q_MOUSE	P01669 mus musculus
28	365.5	62.7	111	1	KV3U_MOUSE	P01673 mus musculus
29	365	62.6	113	1	KV2B_HUMAN	P01615 homo sapien
30	365	62.6	113	1	KV2G_MOUSE	P01631 mus musculus
31	364.5	62.5	108	1	KV1V_HUMAN	P04430 homo sapien
32	364.5	62.5	108	1	KV1Y_HUMAN	P80362 homo sapien
33	364.5	62.5	115	1	KV2A_HUMAN	P01614 homo sapien

ALIGNMENTS

RESULT 1

117	1	KV2E_HUMAN	P06309 homo sapien
108	1	KV1H_HUMAN	P01600 homo sapien
107	1	KV1D_HUMAN	P01596 homo sapien
109	1	KV3G_HUMAN	P04206 homo sapien
131	1	KV3I_MOUSE	P01661 mus musculus
108	1	KV1K_HUMAN	P01603 homo sapien
108	1	KV1Q_HUMAN	P01607 homo sapien
111	1	KV3S_MOUSE	P01671 mus musculus
128	1	KV3K_HUMAN	P06311 homo sapien
113	1	KV2D_HUMAN	P01617 homo sapien
111	1	KV3T_MOUSE	P01672 mus musculus
108	1	KV1F_HUMAN	P01598 homo sapien
108	1	KV1R_HUMAN	P01610 homo sapien
108	1	KV1B_HUMAN	P01594 homo sapien
108	1	KV1L_HUMAN	P01604 homo sapien
111	1	KV3D_MOUSE	P03977 mus musculus
111	1	KV3K_MOUSE	P01663 mus musculus
112	1	KV3G_MOUSE	P01659 mus musculus
108	1	KV1E_HUMAN	P01597 homo sapien
108	1	KV1P_HUMAN	P01608 homo sapien
112	1	KV2C_HUMAN	P01616 homo sapien
112	1	KV3B_MOUSE	P01655 mus musculus
108	1	KV5M_MOUSE	P01646 mus musculus
113	1	KV2E_MOUSE	P03976 mus musculus
108	1	KV1S_HUMAN	P01611 homo sapien
108	1	KV5P_MOUSE	P01649 mus musculus
132	1	KV3F_MOUSE	P01658 mus musculus
108	1	KV5J_MOUSE	P01643 mus musculus
108	1	KV1C_HUMAN	P01595 homo sapien
108	1	KV1Q_HUMAN	P01609 homo sapien
111	1	KV3C_MOUSE	P01656 mus musculus
111	1	KV3E_MOUSE	P01657 mus musculus
129	1	KV1X_HUMAN	P04432 homo sapien
108	1	KV1N_HUMAN	P01606 homo sapien
108	1	KV5S_MOUSE	P01652 mus musculus
113	1	KV2F_MOUSE	P01630 mus musculus
108	1	KV1G_HUMAN	P01599 homo sapien
108	1	KV3V_MOUSE	P01674 mus musculus
108	1	KV5Q_MOUSE	P01650 mus musculus
116	1	KV3J_HUMAN	P04434 homo sapien
128	1	KV5E_MOUSE	P01637 mus musculus
111	1	KV3A_MOUSE	P01634 mus musculus
100	1	KV3C_HUMAN	P01621 homo sapien
108	1	KV1A_HUMAN	P01593 homo sapien
108	1	KV5T_MOUSE	P01653 mus musculus
112	1	KV1U_HUMAN	P01613 homo sapien
117	1	KV1J_HUMAN	P01602 homo sapien
112	1	KV2D_MOUSE	P01629 mus musculus
108	1	KV5D_MOUSE	P01636 mus musculus
108	1	KV5K_MOUSE	P01644 mus musculus
108	1	KV5L_MOUSE	P01645 mus musculus
110	1	KV15_RABIT	P01696 oryctolagus
115	1	KV3I_HUMAN	P04433 homo sapien
108	1	KV5N_MOUSE	P01647 mus musculus
108	1	KV5O_MOUSE	P01648 mus musculus
109	1	KV1T_HUMAN	P01612 homo sapien
113	1	KV2C_MOUSE	P01628 mus musculus
108	1	KV5R_MOUSE	P01651 mus musculus
129	1	KV4A_MOUSE	P01680 mus musculus
108	1	KV06_RABIT	P01687 oryctolagus
112	1	KV2A_MOUSE	P01626 mus musculus
109	1	KV03_RABIT	P01684 oryctolagus
114	1	KV16_RABIT	P01697 oryctolagus
117	1	KV1I_HUMAN	P01601 homo sapien
107	1	KV6A_MOUSE	P01675 mus musculus
108	1	KV1_CANFA	P01618 canis famil
107	1	KV6D_MOUSE	P01678 mus musculus

```
KV4B_HUMAN          STANDARD;          PRT;   133 AA.
ID   KV4B_HUMAN      AC   P06313;
DT   01-JAN-1988 (Rel. 06, Created)
DI   01-JAN-1988 (Rel. 06, Last sequence update)
DI   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-IV region JI precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=6004342; PubMed=50995;
RA   Schneider M., Hilschmann N.;
RT   "The primary structure of a monoclonic immunoglobulin-L-chain of
RL   subgroup IV of the kappa type (Bence-Jones protein Len).";
RL   Hoppe-Seyler's Z. Physiol. Chem. 356:507-557(1975).
RN   [2]
RP   REVISION TO 9.
RA   Salomon A.;
RL   Submitted (AUG-1996) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC   -!- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR   PIR: A01903; KAHULN.
DR   HSSP: P01607; IREI.
DR   InterPro: IPR003006; Ig_MHC.
DR   InterPro: IPR003596; Ig_V.
DR   Pfam: PF00047; Ig; 1.
DR   SMART: SM00406; IG; 1.
KW   Immunoglobulin V region; Bence-Jones protein.
FT   DOMAIN          1 23  FRAMEWORK-1
FT   DOMAIN          24 40  COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN          41 55  FRAMEWORK-2.
FT   DOMAIN          56 62  COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN          63 94  FRAMEWORK-3.
FT   DOMAIN          95 101 COMPLEMENTARITY-DETERMINING-3.
FT   DOMAIN          102 113 FRAMEWORK-4.
FT   DISULFID        23 94  BY SIMILARITY.
FT   NON_TER         114
FT   SEQUENCE        114 AA; 12640 MW; 0647F1D17F236485 CRC64;
SQ
Query Match      89.3%; Score 520.5; DB 1; Length 114;
Best Local Similarity 89.4%; Pred. No. 5.9e-49;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY   1 DIVTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
    |||||
DB   1 DIVTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
QY   61 ESGVPRFSGSGGTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTGLEIK 112
    |||||
DB   61 ESGVPRFSGSGGTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTGLEIK 113

RESULT 3
KV4C_HUMAN          STANDARD;          PRT;   134 AA.
ID   KV4C_HUMAN      AC   P06314;
DT   01-JAN-1988 (Rel. 06, Created)
DI   01-APR-1988 (Rel. 07, Last sequence update)
DI   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-IV region B17 precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86041854; PubMed=2997713;
RA   Marsh P., Mills F., Gould H.;
RT   "Detection of a unique human V kappa IV germline gene by a cloned
RL   cDNA probe.";
RL   Nucleic Acids Res. 13:6531-6544(1985).
RN   [2]
RP   REVISION TO 76.
RA   Marsh P.;
RL   Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
KV4A_HUMAN          STANDARD;          PRT;   114 AA.
ID   KV4A_HUMAN      AC   P01625;
DT   21-JUL-1986 (Rel. 01, Created)
DI   01-OCT-1996 (Rel. 34, Last sequence update)
DI   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-IV region Len.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86041853; PubMed=2997712;
RA   Klobbeck H.G., Bornkamm G.W., Combriato G., Mocikat R., Pohlenz H.D.,
RA   Zachau H.G.;
RT   "Subgroup IV of human immunoglobulin K light chains is encoded by a
RL   single germline gene.";
RL   Nucleic Acids Res. 13:6515-6529(1985).
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
EMBL: Z00022; CAA77317.1; -
PIR: A01904; K4HUJ1.
HSSP: P80362; IWTL.
InterPro: IPR003006; Ig_MHC.
InterPro: IPR003596; Ig_V.
Pfam: PF00047; Ig; 1.
SMART: SM00406; IG; 1.
KW   Immunoglobulin V region; Signal.
FT   SIGNAL          1 20
FT   CHAIN           21 133 IG KAPPA CHAIN V-IV REGION JI.
FT   DOMAIN          21 43  FRAMEWORK-1.
FT   DOMAIN          44 60  COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN          61 75  FRAMEWORK-2.
FT   DOMAIN          76 82  COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN          83 114 FRAMEWORK-3.
FT   DOMAIN          115 122 COMPLEMENTARITY-DETERMINING-3.
FT   DOMAIN          123 132 FRAMEWORK-4.
FT   DISULFID        43 114 BY SIMILARITY.
FT   NON_TER         133
FT   SEQUENCE        133 AA; 14632 MW; 5FB3953066744AF4 CRC64;
SQ
Query Match      89.4%; Score 521; DB 1; Length 133;
Best Local Similarity 89.3%; Pred. No. 6.3e-49;
Matches 100; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY   1 DIVTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
    |||||
DB   21 DIVTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 80
QY   61 ESGVPRFSGSGGTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTGLEIK 112
    |||||
DB   81 ESGVPRFSGSGGTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTGLEIK 132

RESULT 2
KV4A_HUMAN          STANDARD;          PRT;   114 AA.
ID   KV4A_HUMAN      AC   P01625;
DT   21-JUL-1986 (Rel. 01, Created)
DI   01-OCT-1996 (Rel. 34, Last sequence update)
DI   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-IV region Len.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```


modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
-----
DR EMBL: X02990; CAA36733.1; -.
DR PIR: A01905; K4HU17.
DR HSP: P80362; IWTLL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 134 IG KAPPA CHAIN V-IV REGION B17.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 61 75 FRAMEWORK-2.
FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 83 114 FRAMEWORK-3.
FT DOMAIN 115 121 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 122 133 FRAMEWORK-4.
FT NON_TER 134 134 BY SIMILARITY.
SQ SEQUENCE 134 AA; 14966 MW; 6413A22FD0738832 CRC64;

Query Match 86.7%; Score 505.5; DB 1; Length 134;
Best Local Similarity 85.8%; Pred. No. 2.9e-47;
Matches 97; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAWSYQKPGSPKLLIYWASTR 60
Db 21 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAWSYQKPGSPKLLIYWASTR 80
QY 61 ESGVPRFSGSGGTDTLTISVQAEADLAVYCHQYLS-SYTFGGTKLEIK 112
Db 81 ESGVPRFSGSGGTDTLTISVQAEADLAVYCHQYLYPLPTFGGTVK 133

RESULT 4
KV40_HUMAN STANDARD; PRT; 121 AA.
AC P06312;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ig kappa chain V-IV region precursor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86041853; PubMed=2997712;
RA Klobbeck H.G., Bornkamm G.W., Combriato G., Mocikat R., Pohlenz H.D.,
RA Zachau H.G.;
RT "Subgroup IV of human immunoglobulin K light chains is encoded by a
RT single germline gene.";
RL Nucleic Acids Res. 13:6515-6529(1985).
CC -1- MISCELLANEOUS: THERE IS ONLY ONE IG KAPPA V-IV GENE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
FT DOMAIN 30 52 FRAMEWORK-1.
FT DOMAIN 53 63 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 64 78 FRAMEWORK-2.
FT DOMAIN 79 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DOMAIN 118 126 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 127 136 FRAMEWORK-4.
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 14902 MW; 8CDD85113996DIC2 CRC64;

Query Match 66.6%; Score 388.5; DB 1; Length 136;
Best Local Similarity 69.0%; Pred. No. 1.e-34;
Matches 78; Conservative 15; Mismatches 13; Indels 7; Gaps 2;

QY 1 DIVVTQSPDSLAVSLGERTVMTNCKSSQSVLYSSNKNKYLAWYQKQPGQSPKLLIYWASTR 60
Db 30 NIWVTQSPKMSMSVSGERVTLTKASENVV-----TIVSWYQKQPGQSPKLLIYGASNR 83
QY 61 ESGVPRFSGSGGSDTFTLTISVQAEDLAVYCHQ-YLSSYTFGGTGKLEIK 112
L 84 YTGVPDRFTGSGSATDFTLTISVQAEDLADYHCGQGYSPYTFGGTGKLEIK 136

RESULT 6
KV3P_MOUSE STANDARD; PRT; 110 AA.
AC P01668;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 7210.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RA "Rearrangement of genetic information may produce immunoglobulin
RT diversity.";
RL Nature 276:785-790(1978).
DR PIR; D01937; KMS10.
DR HSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 38 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 39 53 FRAMEWORK-2.
FT DOMAIN 54 60 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 61 92 FRAMEWORK-3.
FT DOMAIN 93 100 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 101 110 FRAMEWORK-4.
FT DISULFID 23 92 BY SIMILARITY.
FT NON_TER 110 110
SQ SEQUENCE 110 AA; 11950 MW; 69F1A5CE886B1249 CRC64;

Query Match 66.4%; Score 387; DB 1; Length 110;
Best Local Similarity 67.9%; Pred. No. 1.e-34;
Matches 76; Conservative 12; Mismatches 22; Indels 2; Gaps 1;

QY 1 DIVVTQSPDSLAVSLGERTVMTNCKSSQSVLYSSNKNKYLAWYQKQPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGORATISCKASQSDYDGD--SYMNWYQKQPGQSPKLLIYAASNL 58
QY 61 ESGVPRFSGSGGSDTFTLTISVQAEDLAVYCHQYLSSTYTFGGTGKLEIK 112
L 84 YTGVPDRFTGSGSATDFTLTISVQAEDLADYHCGQGYSPYTFGGTGKLEIK 136
Db 59 ESGIPARFSGSGGSDTFTLTINHPVEEDAATYTCQSDPDWTFGSGTGKLEIK 110

RESULT 7
KV3L_MOUSE STANDARD; PRT; 111 AA.
AC P01664;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region CBPC 101.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
MEDLINE=79012520; PubMed=99744;
RA McKean D.J., Bell M., Potter M.;
RA "Mechanisms of antibody diversity: multiple genes encode structurally
RT related mouse kappa variable regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.
DR PIR; A01936; KVMSC1.
DR HSP; P80362; IWTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 38 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 39 53 FRAMEWORK-2.
FT DOMAIN 54 60 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 61 92 FRAMEWORK-3.
FT DOMAIN 93 101 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 102 111 FRAMEWORK-4.
FT DISULFID 23 92 BY SIMILARITY.
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 11964 MW; E2BIAD98AD965962 CRC64;

Query Match 66.0%; Score 384.5; DB 1; Length 111;
Best Local Similarity 69.0%; Pred. No. 2.le-34;
Matches 78; Conservative 11; Mismatches 21; Indels 3; Gaps 2;

QY 1 DIVVTQSPDSLAVSLGERTVMTNCKSSQSVLYSSNKNKYLAWYQKQPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGORATISCKASQSDYTG--ESYMNWYQKQPGQSPKLLIYAASNL 58
QY 61 ESGVPRFSGSGGSDTFTLTISVQAEDLAVYCHQYLSSTYTFGGTGKLEIK 112
L 84 YTGVPDRFTGSGSATDFTLTINHPVEEDAATYTCQSDPDWTFGSGTGKLEIK 111
Db 59 ESGIPARFSGSGGSDTFTLTINHPVEEDAATYTCQSDPDWTFGSGTGKLEIK 111

RESULT 8
KV3L_HUMAN STANDARD; PRT; 129 AA.
AC P18135;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region HAH precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=88171307; PubMed=3127527;
RA Kipps T.J., Tomhave E., Chen P.P., Carson D.A.;
RA "Antibody-associated kappa light chain variable region gene
RT expressed in chronic lymphocytic leukemia with little or no somatic
RT mutation. Implications for etiology and immunotherapy.";
RL J. Exp. Med. 167:840-852(1988).
CC -1- DISEASE: THE PROTEIN IS ONE OF THE SURFACE IMMUNOGLOBULIN M
CC AUTOANTIBODIES EXPRESSED IN PATIENTS WITH CHRONIC LYMPHOCYTIC
```

```
CC DR LEUKEMIA.
DR PIR; P00022; K3HUHA.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 129 IG KAPPA CHAIN V-III REGION HAH.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 21 43 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 44 55 FRAMEWORK-2.
FT DOMAIN 56 70 FRAMEWORK-3.
FT DOMAIN 71 77 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 78 109 FRAMEWORK-3.
FT DOMAIN 110 118 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 119 129 JK1 SEGMENT.
FT DISULFID 43 109 BY SIMILARITY.
FT NON_TER 129 129
SQ SEQUENCE 129 AA; 14073 MW; D3C55292772774D0 CRC64;

Query Match 65.4%; Score 381; DB 1; Length 129;
Best Local Similarity 65.5%; Pred. No. 6e-34;
Matches 74; Conservative 18; Mismatches 15; Indels 6; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKQKPSKLLIYMASTR 60
DB 21 EIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAWYQKQKQAPRLLIYGASSR 75

QY 61 ESGVDPFRSGSGGTDTFTLTSSVQAEDLAVYVCHQY-LSSYFGGQTKLEIK 112
DB 76 ATGIPDRFSGSGGTDTFTLTISRLEPDAFYVYCOQYGSPTFGQGTKEIK 128

RESULT 9
KV3M_HUMAN
ID KV3M_HUMAN STANDARD; PRT; 129 AA.
AC P18136;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region HIC precursor.
OS Homo sapiens. (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=88171307; PubMed=3127527;
Kipps T.J., Tomhave E., Chen P.P., Carson D.A.;
"Autoantibody-associated kappa light chain variable region gene
expressed in chronic lymphocytic leukemia with little or no somatic
mutation. Implications for etiology and immunotherapy.";
J. Exp. Med. 167:840-852(1988).
CC -1- DISEASE: THE PROTEIN IS ONE OF THE SURFACE IMMUNOGLOBULIN M
ANTOANTIBODIES EXPRESSED IN PATIENTS WITH CHRONIC LYMPHOCYTIC
LEUKEMIA.
DR PIR; P00021; K3HUH1.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 129 IG KAPPA CHAIN V-III REGION HIC.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 55 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 56 70 FRAMEWORK-2.
FT DOMAIN 71 77 FRAMEWORK-3.
FT DOMAIN 78 109 FRAMEWORK-3.
FT DOMAIN 110 118 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 119 129 JK1 SEGMENT.

FT DISULFID 43 109 BY SIMILARITY.
FT NON_TER 129 129
SQ SEQUENCE 129 AA; 14070 MW; 7395528EA2BB74D6 CRC64;

Query Match 65.4%; Score 381; DB 1; Length 129;
Best Local Similarity 65.5%; Pred. No. 6e-34;
Matches 74; Conservative 19; Mismatches 14; Indels 6; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKQKPSKLLIYMASTR 60
DB 21 EIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAWYQKQKQAPRLLIYGASSR 75

QY 61 ESGVDPFRSGSGGTDTFTLTSSVQAEDLAVYVCHQY-LSSYFGGQTKLEIK 112
DB 76 ATGIPDRFSGSGGTDTFTLTISRLEPDAFYVYCOQYGSPTFGQGTKEIK 128

RESULT 10
KV5A_MOUSE
ID KV5A_MOUSE STANDARD; PRT; 149 AA.
AC P01633;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region MPC11 precursor.
OS Mus musculus. (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-71 FROM N.A.
MEDLINE=83001944; PubMed=6288267;
Kelley D.E., Coleclough C., Perry R.P.;
"Functional significance and evolutionary development of the
5'-terminal regions of immunoglobulin variable-region genes.";
Cell 29:681-689(1982).
CC -1- MISCELLANEOUS: THE MATURE CHAIN HAS 12 ADDITIONAL RESIDUES AT ITS
AMINO END, DUE TO A TANDEM DUPLICATION OF 36 NUCLEOTIDES AFTER THE
CODON FOR RESIDUE 36. RESIDUE 42 CORRESPONDS TO THE AMINO-TERMINAL
RESIDUE OF TYPICAL KAPPA CHAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; J00561; AAA38776.1; -.
DR PIR; A01916; KYMS11.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal; Repeat.
FT SIGNAL 1 29
FT CHAIN 30 149 IG KAPPA CHAIN V-V REGION MPC11.
```



```

DR EMBL; U29423; AAC00033.1; .
DR PIR; A01915; KVM57A.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IG; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 40
FT DOMAIN 41 55
FT DOMAIN 56 62
FT DOMAIN 63 94
FT DOMAIN 95 103
FT DOMAIN 104 113
FT DOMAIN 114 114
FT DISULFID 23 94
FT NON_TER 114 114
SQ SEQUENCE 114 AA; 12717 MW; 32008EC8E9DBE67B CRC64;

Query Match 64.4%; Score 375.5; DB 1; Length 114;
Best Local Similarity 69.3%; Pred. No. 2e-33;
Matches 79; Conservative 12; Mismatches 20; Indels 3; Gaps 3;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQK-NYLAWYQKPGQSPKLLIYWAST 59
DB 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQK-NYLAWYQKPGQSPKLLIYGASN 59
QY 60 RESGVPDRFSGSGSGTDFTLTISVQAEDLAVYVCHQYLS-SYTFGGGKLEIK 112
DB 60 RYTGVPDRFSGSGSGTDFTLTISVQAEDLAVYVCHQYLS-SYTFGGGKLEIK 113

RESULT 14
KVIM_HUMAN STANDARD; PRT; 108 AA.
AC P01605;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Lay.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77038198; PubMed=824717;
RA Capra J.D., Klapper D.G.;
RT "Complete amino acid sequence of the variable domains of two human
RT IgM anti-gamma globulins (Lay/Pom) with shared idiotypic
RT specificities."
RL Scand. J. Immunol. 5:677-684(1976).
CC -1- MISCELLANEOUS: THE SECOND AND THIRD HYPERVARIABLE REGIONS OF THIS
CC CHAIN ARE IDENTICAL WITH THOSE OF THE HUMAN POM V-III KAPPA CHAIN,
CC WITH WHICH IT SHARES CERTAIN IDIOTYPIC DETERMINANTS.
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR; A01871; KIHULY.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IG; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 97
FT DOMAIN 98 107
FT DOMAIN 108 108
FT DISULFID 23 88
FT NON_TER 108 108

Query Match 64.2%; Score 374.5; DB 1; Length 108;
Best Local Similarity 66.4%; Pred. No. 2.4e-33;
Matches 75; Conservative 14; Mismatches 17; Indels 7; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQK-NYLAWYQKPGQSPKLLIYWASTR 60
DB 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQK-NYLAWYQKPGQSPKLLIYGASTR 54
QY 61 ESCVPRFRFSGSGSGTDFTLTISVQAEDLAVYVCHQYLS-SYTFGGGKLEIK 112
DB 55 EAGVPSRFRSGSGSGTDFTLTISVQAEDLAVYVCHQYLS-SYTFGGGKLEIK 107

RESULT 15
KV3E_HUMAN STANDARD; PRT; 109 AA.
ID KV3E_HUMAN
AC P01623;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region WOL.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=82046598; PubMed=6794615;
RA Andrews D.W., Capra J.D.;
RT "Amino acid sequence of the variable regions of light chains from two
RT idiotypically cross-reactive human Igm anti-gamma-globulins of the Wa
RT group."
RL Biochemistry 20:5816-5822(1981).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR; A01896; K3HULW.
DR HSSP; P80362; IWLW.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IG; 1.
KW Immunoglobulin V region.
FT DISULFID 23 89
FT NON_TER 109 109
FT NON_TER 109 109
BY SIMILARITY.
SQ SEQUENCE 109 AA; 11746 MW; 566C115E689CBEE CRC64;

Query Match 64.2%; Score 374; DB 1; Length 109;
Best Local Similarity 64.6%; Pred. No. 2.8e-33;
Matches 73; Conservative 17; Mismatches 17; Indels 6; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQK-NYLAWYQKPGQSPKLLIYWASTR 60
DB 1 EIVLTQSPGTLSPGGERATLSCRASQSV-----SSGYLGWYQKPGQAPRLIYGASSR 55
QY 61 ESCVPRFRFSGSGSGTDFTLTISVQAEDLAVYVCHQYLS-SYTFGGGKLEIK 112
DB 56 ATGIDPRFRFSGSGSGTDFTLTISVQAEDLAVYVCHQYLS-SYTFGGGKLEIK 108

RESULT 16
KV3M_MOUSE STANDARD; PRT; 111 AA.
ID KV3M_MOUSE
AC P01665;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 7043.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN  SEQUENCE.
RX  MEDLINE=79073152; PubMed=103003;
RA  Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT  "Rearrangement of genetic information may produce immunoglobulin
RL  diversity.";
DR  Nature 276:785-790(1978).
DR  PIR: A01937; KVM543.
DR  HSSP; P80362; LWTL.
DR  InterPro; IPR003006; Ig_MHC.
DR  InterPro; IPR003596; Ig_v.
DR  Pfam; PF00047; Ig; 1.
DR  SMART; SM00406; IGV; 1.
KW  Immunoglobulin v region.
FT  DOMAIN 1 23
FT  DOMAIN 24 38
FT  DOMAIN 39 53
FT  DOMAIN 54 60
FT  DOMAIN 61 92
FT  DOMAIN 93 101
FT  DOMAIN 102 111
FT  DISULFID 23 92
FT  NON_TER 111 111
SQ  SEQUENCE 111 AA; 12002 MW; 7A5FCB586C306D29 CRC64;

Query Match
Best Local Similarity 64.1%; Score 373.5; DB 1; Length 111;
Matches 76; Conservative 11; Mismatches 23; Indels 3; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGORATISCKASQSDYDGD--SYMNWYQKPGQPKLLIYAASNL 58

QY 61 ESGVPRFSGSGGTDTLTITSSVQAEADLAVYCHQYLSS-YTFGGGTKLEIK 112
Db 59 ESGIPARFSGSGGTDTLTINHPVEEDAATYYCQSNEDPFTFGGSKLEIK 111

RESULT 17
KV3O_MOUSE STANDARD; PRT; 111 AA.
AC P01667;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 6308.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN  SEQUENCE.
RX  MEDLINE=79073152; PubMed=103003;
RA  Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT  "Rearrangement of genetic information may produce immunoglobulin
RL  diversity.";
DR  Nature 276:785-790(1978).
DR  PIR: A01937; KVM508.
DR  HSSP; P80362; LWTL.
DR  InterPro; IPR003006; Ig_MHC.
DR  InterPro; IPR003596; Ig_v.
DR  Pfam; PF00047; Ig; 1.
DR  SMART; SM00406; IGV; 1.
KW  Immunoglobulin v region.
FT  DOMAIN 1 23
FT  DOMAIN 24 38
FT  DOMAIN 39 53
FT  DOMAIN 54 60
FT  DOMAIN 61 92
FT  DOMAIN 93 101
FT  DOMAIN 102 111
FT  DISULFID 23 92
FT  NON_TER 111 111
SQ  SEQUENCE.
RX  MEDLINE=79073152; PubMed=103003;
RA  Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT  "Rearrangement of genetic information may produce immunoglobulin
RL  diversity.";
DR  Nature 276:785-790(1978).
DR  PIR: A01937; KVM537.
DR  HSSP; P01679; 2FBJ.
DR  InterPro; IPR003006; Ig_MHC.
DR  InterPro; IPR003596; Ig_v.
DR  Pfam; PF00047; Ig; 1.
DR  SMART; SM00406; IGV; 1.
KW  Immunoglobulin v region.
FT  DOMAIN 1 23
FT  DOMAIN 24 38
FT  DOMAIN 39 53
FT  DOMAIN 54 60
FT  DOMAIN 61 92
FT  DOMAIN 93 101
FT  DOMAIN 102 111
FT  DISULFID 23 92
FT  NON_TER 111 111
SQ  SEQUENCE 111 AA; 12099 MW; EC46C9D259213BE4 CRC64;

Query Match
Best Local Similarity 63.9%; Score 372.5; DB 1; Length 111;
Matches 75; Conservative 15; Mismatches 20; Indels 3; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGORATISCKASESV--DSYGNSEFMHWYQKPGQPKLLIYASNL 58

QY 61 ESGVPRFSGSGGTDTLTITSSVQAEADLAVYCHQYLSS-YTFGGGSKLEIK 112
Db 59 ESGIPARFSGSGGTDTLTINHPVEADVDVATYYCQSNEDPFTFGGSKLEIK 111
```

```

SQ  SEQUENCE 111 AA; 12071 MW; 7A4ADE4D6C256D29 CRC64;

Query Match
Best Local Similarity 64.1%; Score 373.5; DB 1; Length 111;
Matches 76; Conservative 11; Mismatches 23; Indels 3; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGORATISCKASQSDYDGD--SYMNWYQKPGQPKLLIYTASNL 58

QY 61 ESGVPRFSGSGGTDTLTITSSVQAEADLAVYCHQYLSS-YTFGGGTKLEIK 112
Db 59 ESGIPARFSGSGGTDTLTINHPVEEDAATYYCQSNEDPFTFGGSKLEIK 111

RESULT 18
KV3H_MOUSE STANDARD; PRT; 111 AA.
AC P01660;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 3741/TEPC 111.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN  SEQUENCE (PC 3741).
RX  MEDLINE=79073152; PubMed=103003;
RA  Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT  "Rearrangement of genetic information may produce immunoglobulin
RL  diversity.";
RL  Nature 276:785-790(1978).
RN  SEQUENCE (TEPC 111).
RX  MEDLINE=79073152; PubMed=99744;
RA  McKean D.J., Bell M., Potter M.;
RT  "Mechanisms of antibody diversity: multiple genes encode structurally
RL  related mouse kappa variable regions.";
RL  Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).
CC  -1- MISCELLANEOUS: THE PC 3741 AND TEPC 111 SEQUENCES ARE IDENTICAL.
DR  PIR; A01934; KVM537.
DR  HSSP; P01679; 2FBJ.
DR  InterPro; IPR003006; Ig_MHC.
DR  InterPro; IPR003596; Ig_v.
DR  Pfam; PF00047; Ig; 1.
DR  SMART; SM00406; IGV; 1.
KW  Immunoglobulin v region.
FT  DOMAIN 1 23
FT  DOMAIN 24 38
FT  DOMAIN 39 53
FT  DOMAIN 54 60
FT  DOMAIN 61 92
FT  DOMAIN 93 101
FT  DOMAIN 102 111
FT  DISULFID 23 92
FT  NON_TER 111 111
SQ  SEQUENCE 111 AA; 12099 MW; EC46C9D259213BE4 CRC64;

Query Match
Best Local Similarity 63.9%; Score 372.5; DB 1; Length 111;
Matches 75; Conservative 15; Mismatches 20; Indels 3; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGORATISCKASESV--DSYGNSEFMHWYQKPGQPKLLIYASNL 58

QY 61 ESGVPRFSGSGGTDTLTITSSVQAEADLAVYCHQYLSS-YTFGGGSKLEIK 112
Db 59 ESGIPARFSGSGGTDTLTINHPVEADVDVATYYCQSNEDPFTFGGSKLEIK 111
```

```
RESULT 19
KV3H_HUMAN STANDARD; PRT; 129 AA.
AC P04207;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region CLL precursor (Rheumatoid factor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86177570; PubMed=3083417;
RA Jirik F.R., Sorge J., Fong S., Heitzmann J.G., Curd J.G., Chen P.P.,
RA Goldfien R., Carson D.A.;
RT "Cloning and sequence determination of a human rheumatoid factor
light-chain gene."
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL: M12740; AAA58992.1;
DR PIR: A01898; K3HUC1.
DR HSSP: P80362; IWT1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 129 IG KAPPA CHAIN V-III REGION CLL.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 69 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 108 FRAMEWORK-3.
FT DOMAIN 109 118 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 119 129 JK1 SEGMENT.
FT DISULFID 43 108 BY SIMILARITY.
FT NON_TER 129 129
FT SEQUENCE 129 AA; 14275 MW; 5C13B411BE60CC14 CRC64;
Query Match 63.8%; Score 372; DB 1; Length 129;
Best Local Similarity 64.9%; Pred. No. 5.5e-33;
Matches 74; Conservative 17; Mismatches 15; Indels 8; Gaps 2;
QY 1 DIVMTQSPDLSVSLGGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
DB 21 EIVMTQSPATLSVSGERATLSKASQSV-----SNLAHYQKPGQPPRLIYGASTR 74
QY 61 ESGVDPFRFSGSGGTDFLTITSSVQAEADLVVYCHQVLS--SYTFGGTKLEIK 112
DB 75 ATGIPARFSGSGGTDFLTITSLRQSEDFAVYVCCQYNWPPWTFGGTRVEIK 128
RESULT 20
KV3N_MOUSE STANDARD; PRT; 111 AA.
AC P01666;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 7183.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
diversity."
RL Nature 276:785-790(1978).
DR PIR: B01937; KVM583.
DR HSSP: P01679; 2FBJ.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 38 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 39 53 FRAMEWORK-2.
FT DOMAIN 54 60 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 61 92 FRAMEWORK-3.
FT DOMAIN 93 101 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 102 111 FRAMEWORK-4.
FT DISULFID 23 92 BY SIMILARITY.
FT NON_TER 111 111
FT SEQUENCE 111 AA; 11952 MW; 2058BB50CE306D31 CRC64;
Query Match 63.0%; Score 367.5; DB 1; Length 111;
Best Local Similarity 66.4%; Pred. No. 1.4e-32;
Matches 75; Conservative 11; Mismatches 24; Indels 3; Gaps 2;
QY 1 DIVMTQSPDLSVSLGGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
DB 1 DIVLTQSPASLSLQGRATISCKASQSVYDGD--SYMNNYQKPGQPPKLLIYAASNL 58
QY 61 ESGVDPFRFSGSGGTDFLTITSSVQAEADLVVYCHQVLS--YTFGGTKLEIK 112
DB 59 ESGIPARFSGSGGTDFLTINHPVEEADAATYVCCQSNEDPLTFGAGTKLEIK 111
RESULT 21
KV3R_MOUSE STANDARD; PRT; 111 AA.
AC P01670;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 6684.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
diversity."
RL Nature 276:785-790(1978).
DR PIR: A01938; KVM584.
DR HSSP: P80362; IWT1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 38 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 39 53 FRAMEWORK-2.
FT DOMAIN 54 60 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 61 92 FRAMEWORK-3.
FT DOMAIN 93 101 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 102 111 FRAMEWORK-4.
```

```
FT DISULFID 23 92 BY SIMILARITY.
FT NON_TER 111
SQ SEQUENCE 111 AA; 12039 MW; 1E46988341858526 CRC64;

Query Match
Best Local Similarity 63.0%; Score 367.5; DB 1; Length 111;
Matches 76; Conservative 12; Mismatches 22; Indels 3; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMNCSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASIALSLGQRTATSCRSKSV--STGYSYMHVYQKPGQPPKLLIYASNL 58

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQVLSSTYFGGGTKLEIK 112
Db 59 ESGVPARESGSGGTDFLTINHPVEEDAAYVYCHSRELPRRTGGGTKLEIK 111

RESULT 22
KV1W_HUMAN STANDARD; PRT; 129 AA.
ID P04431.
RA 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Walker precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85014148; PubMed=6091049;
RA Klobeck H.G., Combrato G., Zachau H.G.;
RT "Immunoglobulin genes of the kappa light chain type from two human
RL lymphoid cell lines are closely related.";
RL Nucleic Acids Res. 12:6995-7006(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X00965; CAA25477.1; ALT_TERM.
DR PIR; A01883; K1HUWK.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 22
FT CHAIN 23 129 IG KAPPA CHAIN V-I REGION WALKER.
FT DOMAIN 23 45 FRAMEWORK-1.
FT DOMAIN 23 45 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 46 56 FRAMEWORK-2.
FT DOMAIN 57 71 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 72 78 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 79 110 FRAMEWORK-3.
FT DOMAIN 111 119 COMPLEMENTARITY-DETERMINING-4.
FT DOMAIN 120 129 FRAMEWORK-4.
FT DISULFID 45 110 BY SIMILARITY.
FT NON_TER 129 129
SQ SEQUENCE 129 AA; 14069 MW; F941FA07D4AFC2F9 CRC64;

Query Match
Best Local Similarity 63.0%; Score 367.5; DB 1; Length 129;
Matches 74; Conservative 14; Mismatches 18; Indels 7; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMNCSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 23 DIOMTQSPSSLSASVGRVITITCRASQSI-----SNLYNMYQKPGKAPKLLIYAASL 76
```

```
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQVLSSTYFGGGTKLEIK 112
Db 77 QSGVTSRFRSGSGGTDFLTITSSLPEDSATYVCOQSYSTLTITFGQGRLEIK 129

RESULT 23
KV3A_HUMAN STANDARD; PRT; 108 AA.
ID KV3A_HUMAN
AC P01619;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region B6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RA Milstein C.;
RT "The basic sequences of immunoglobulin kappa chains: sequence studies
of Bence Jones proteins Rad, Fr4 and B6.";
RL FEBS Lett. 2:301-304(1969).
CC -!- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR PIR; A01891; K3HUB6.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DISULFID 23 89 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11635 MW; 8BC14FF07A419E3D CRC64;

Query Match
Best Local Similarity 63.0%; Score 367; DB 1; Length 108;
Matches 69; Conservative 22; Mismatches 16; Indels 6; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMNCSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 ZIVLTZSPGTLSLSPGZRAALSCRASQSL-----SGNYLAHYQKPGQAPLLMYGVSSR 55

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQVLSSTYFGGGTKLEIK 112
Db 56 ATGIPDRFSGSGGTDFLTITSLRLZPEDFAVYVCOQYQYGVSSPFTFGQSKLEIK 108

RESULT 24
KV3F_HUMAN STANDARD; PRT; 109 AA.
ID KV3F_HUMAN
AC P01624;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig kappa chain V-III region POM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76276460; PubMed=60899;
RA Klapper D.G., Capra J.D.;
RT "The amino acid sequence of the variable regions of the light chains
from two idiotypically cross reactive IgM anti-gamma globulins.";
RL Ann. Immunol. (Paris) 127C:261-271(1976).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR; A01897; K3HUPM.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
```


DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig: 1.
DR SMART: SM00406; IGV: 1.
KW Immunoglobulin V region.
FT DISULFID 23 89
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 11922 MW; 62821DDC6A8ABA86 CRC64;

BY SIMILARITY.

Query Match 63.0%; Score 367; DB 1; Length 109;
Best Local Similarity 64.6%; Pred. NO. 1.5e-32;
Matches 73; Conservative 18; Mismatches 16; Indels 6; Gaps 2;

QY 1 DIVVTQSPDSLAVSLGERVTWNCKSSQSVLYSSNOKNYLAWYQOKPGOSP KLLIYWASTR 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 EIVMTQSPVTLVSFGERATISCRASQSI-----SNSYLAWYQOKPGSPRLIIYGASTR 55

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLS-SYTFGGGKLEIK 112
:||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 56 ATGIPARFSGSGGTFTLTSSLOSEDFAVYCOQYNNWPTFGQGRVEIK 108

RESULT 25

KV3J_MOUSE
ID KV3J_MOUSE STANDARD; PRT: 111 AA.
AC P01662;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region ABPC 22/PC 9245.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE (ABPC 22).
RX MEDLINE=79012520; PubMed=99744;
RA McKean D.J., Bell M., Potter M.;
RT "Mechanisms of antibody diversity: multiple genes encode structurally
related mouse kappa variable regions."
RL Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).
RN [2]
RP SEQUENCE (PC 9245).
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gattalman L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
diversity."
RL Nature 276:785-790(1978).
-1- MISCELLANEOUS: THE ABPC22 AND PC9241 SEQUENCES ARE IDENTICAL.
PIR: A01935; KVM5K6.
HSP: P01679; 2FBJ.

DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig: 1.
DR SMART: SM00406; IGV: 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38
FT DOMAIN 39 53
FT DOMAIN 54 60
FT DOMAIN 61 92
FT DOMAIN 93 101
FT DOMAIN 102 111
FT DISULFID 23 92
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12041 MW; D7DF0609303453CE CRC64;

Query Match 62.9%; Score 366.5; DB 1; Length 111;
Best Local Similarity 66.4%; Pred. NO. 1.8e-32;
Matches 75; Conservative 13; Mismatches 22; Indels 3; Gaps 2;

QY 1 DIVVTQSPDSLAVSLGERVTWNCKSSQSVLYSSNOKNYLAWYQOKPGOSP KLLIYWASTR 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 NIVLTQSPASLVSLGQRATISCRASESV--DSYGNSEFMHWYQOKPGPPKLLIYASN 58

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVVYCHQYLS-SYTFGGGKLEIK 112
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 59 ESGVPAREFSGSGRTDFTLTIDPVEADDAATYTCQNNEDPYTFGGGKLEIK 111

Search completed: June 23, 2003, 14:02:09
Job time : 5.17021 secs

